THE RELATION BETWEEN FREQUENCY OF RESPONSE AND RETENTION, ON A PAIRED ASSOCIATES TASK

by

EUGENE M. BERQUAM

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL
OF THE UNIVERSITY OF FLORIDA
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1981

© Copyright 1981

by

Eugene M. Berquam

ACKNOWLEDGMENTS

There are many who share in the accomplishments of this manuscript. To my mother and my entire family goes a sincere thank you for your years of tolerating and moderating the behaviors of a dreamer.

The professors who served on my doctoral committee will long be remembered. To Ron Nutter, who informed me (many times) of the distinction between benefit cost and cost effectiveness; to Cecil Mercer, who challenged me to attempt to validate the "fluency" issue; to Bob Algozzine, who, clearly, helped improve my skills with the computer and with the backhand return of service; to Hank Pennypacker, who knows the worth of the science of behavior and the glory of salmo trutta; thank you all for your help and guidance.

For Bill Wolking, chairman of my committee, I will carry a lifelong appreciation. The term "major advisor" has never been more appropriately used than in describing his influence on my career. Bill, the completion of this paper is due largely to your encouragement, guidance, understanding, and friendship. Thank you seems hardly enough.

A few words about the real teachers of educational knowledge--Kids. They have the answers. We need to ask the right questions.

One last message of appreciation goes to those who have pioneered in the development of behavior analysis and precision teaching. I have not been associated with a more professional, competent, and sharing group. Thanks to all of you.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS iii

LIST OF TABLES. vii

ABSTRACT. viii

Chapter

I. INTRODUCTION 1

Rationale 1
Statement of Problem. 4
Limitations 5
Delimitations 5

II. REVIEW OF RELATED LITERATURE 6

Scope of Review 6
Sources Used for Review 7
Chronological Overview. 8
Subjects, Intervals, and Related Issues 10
Quantification of Degree of Learning. 12
Individual Studies. 18
Verbal Learning Studies. 19
Motor Learning Studies 27
Summary 33

III. METHOD 37

Experimental Questions. 37
Independent Variables 38
Dependent Variable. 38
Units of Measurement. 38
Subjects. 39
Experimental Task 39
Procedure 40
Data Analysis 41

Chapter Page

IV. RESULTS 44

Descriptive Data on Subjects. 44
General Analysis Procedures 45
Diagram of Retention Measures 46
Question 1 51
Summary of Question 1 Results 56
Question 2. 56
Question 2B 59
Summary of Question 2 Results 61
Comparison of Retention Before and After Fluency Training 63

V. DISCUSSION. 64

Question 1. 64
Question 2, Matching Comparisons. 68
Question 2, Part A. 69
Question 2, Part B. 71
"Cost" of Training Procedures 72
Educational Applications. 72
Implications for Further Research 73

APPENDICES

A AC-4 BEHAVIOR CHART. 77

B EXAMPLES OF BASIC RELATIONS BETWEEN TWO DATA SETS. . . 79

C EXAMPLE OF FREQUENCY DISTRIBUTION ON AC-4. 81

D EXAMPLE OF CELERATION STACK ON AC-4. 83

E EXAMPLE OF STRAT ANALYSIS. 85

F RE-ANALYSIS OF DATA FROM LUH (1922). 87

G RE-ANALYSIS OF DATA FROM KRUEGER (1929). 89

H RE-ANALYSIS OF DATA FROM UNDERWOOD (1954). 91

I RE-ANALYSIS OF DATA FROM HELLYER (1962). 93

J RE-ANALYSIS OF DATA FROM UNDERWOOD & KEPPEL (1963) . . 95

K RE-ANALYSIS OF DATA FROM CLEMENT (1978). 96

L LITERATURE REVIEW SUMMARY TABLES 99

M EXPERIMENTAL TASK PRACTICE SHEETS. 107

APPENDICES Page

N CORRELATION COEFFICIENTS BETWEEN A SHORT TERM MEMORY TASK AND SCHOOL ACHIEVEMENT (Berquam, 1981) . . .110

O EXPERIMENTAL DESIGN COMPONENTS 112

P MATCHING COMPARISONS 114

Q CORRELATION COEFFICIENTS BETWEEN PERFORMANCE ON THE EXPERIMENTAL TASK AND SCHOOL ACHIEVEMENT. 116

R CHARTED DATA, QUESTION 1 118

S CHARTED DATA, QUESTION 2, FREQUENCY DISTRIBUTIONS. . .122

T CHARTED DATA, QUESTION 2, EXPERIMENTAL-CONTROL COMPARISONS 127

U CHARTED DATA, QUESTION 2, EXPERIMENTAL-CONTROL COMPARISONS, MATCHED ON P5 132

V CHARTED DATA, QUESTION 2, COMPARISON OF TIME ON TASK AND TOTAL NUMBER OF RESPONSES 135

BIBLIOGRAPHY. 136

BIOGRAPHICAL SKETCH 141

LIST OF TABLES

Table		Page
1.	Experimental Group Data, Sorted by P5	48
2.	Control Group Data, Sorted by P5.	49
3.	Correlations Between Various Performance Measures	51
4.	Comparison of Performance on P6, Sorted by Rank on P5	52
5.	Comparison of Performance on P10, Sorted by Rank on P5.	53
6.	Comparison of Performance on P11, Sorted by Rank on P5.	53
7.	Summary Matrix of Contrast Results for Analysis of Variance on P6, P10, P11.	54
8.	Summary Matrix for STRAT Analysis on Retention Measures	55
9.	Medians and Ranges for Distributions on P5, P6, P10, and P11, Sorted by Group.	57
10.	Comparisons of Group Means on Performance Measures.	58
11.	Summary Matrix of STRAT Analysis Comparing Experimental and Control Group Scores, Sorted by P5.	58
12.	Matched Medians on Measure P5	59
13.	Summary Matrix of Comparisons Between Control and Experimental Groups, Matched on P5.	60
14.	Median Scores for Time on Task and Total Responses.	62

Abstract of Dissertation Presented to the Graduate Council
of the University of Florida in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

THE RELATION BETWEEN FREQUENCY OF RESPONSE AND RETENTION, ON A PAIRED ASSOCIATES TASK

By

Eugene M. Berquam

June, 1981

Chairman: William D. Wolking
Major Department: Special Education

Retention has consistently been shown to be strongly related to previous performance or learning level. Thirty-seven studies were reviewed in which degree of learning or amount of practice was systematically varied. Twenty-nine studies showed greater retention for the higher levels of learning. It appeared that a ceiling effect limited results in the other eight studies, due to the specific types of measurements used. Eleven studies provided specific measures of performance improvement. All of these studies showed greater retention to be associated with greater performance.

Two questions were generated from this review. The first concerns the nature of the retention slopes, and whether the percentage or proportion of retention changes with degree of previous performance. The second question concerns the difference between effects due to extra practice and those due to an actual improvement in performance. The

present study used a paired associates task, requiring subjects to write a one-digit number in response to the presentation of a three-letter nonsense syllable. Two classrooms of third graders served as subjects, with a total n of 34.

A two-week retention interval followed the initial five-day learning period. This was followed by a five-day relearning phase, and then another two-week retention interval. Results related to the percentage of retention question indicate that in nine of twelve contrasts, higher initial performance levels are associated with a higher percentage of subsequent retention.

To investigate the difference between extra practice and greater performance, one of the two groups received practice specifically designed to increase the rate of correct responses. This was termed fluency training. The control group received this training during the relearning phase, but received only unpaced practice during initial learning. Comparison of results between the types of training indicates that the fluency training produced greater performance and greater retention than did the non-paced practice, even though the unpaced practice provided more time on task and a greater number of total responses during the learning phase. When the control group received the fluency training during the relearning phase, subsequent retention was greater for this group than when it received only the unpaced practice. The fluency training procedure would seem to be of value in education as an efficient method of improving performance and retention.

CHAPTER I

INTRODUCTION

Mastery and retention of learned material are primary goals of education. Yet, the questions of what constitutes mastery of a specific skill, and how this mastery affects retention have not been clearly answered. Indeed, often they have not been asked. Since the term "mastery" carries strong connotations of an absolute, percentage-correct type of score, it will not be used here. Rather, "fluency," which implies high rate, effortless, error free performance (Haring, Lovitt, Eaton and Hansen, 1978) will be used. It is the intent of this study to examine the relation between frequency of response and retention of a learned skill.

The term "degree of learning" is often used in the literature. Degree of learning can be used as a general term to indicate a wide range in performance. Fluency would be towards the top of that range.

Rationale

Evidence that education is less than universally successful in teaching skills is not difficult to find. A recent Ford Foundation study reported that up to 64 million adults may be illiterate, and that half the adults in nine southern states have not completed high school (Hunter & Harman, 1979). The National Advisory Committee on Dyslexia

estimated that 15% of the school population may have a severe reading problem (Kline, 1972).

In an attempt to eliminate problems such as these, there is a growing trend among states to pass "minimum competency" legislation for education. Nearly 40 states have passed a law of this type (Frahm & Covington, 1979). The Florida state legislature recently enacted the Primary Education Program, or PREP Bill.* This law requires passage of a third-grade state criterion test before students can be promoted into grade four. Failure to pass the test requires that students be retained for one year, unless the teacher can certify that deficient skills have been remediated.

Financial implications of programs such as the PREP Bill are enormous. A district the size of Orange County, in Central Florida, may enroll 7,000 third graders. Using the NACD incidence figure of 15%, a district this size could have 1,000 third graders with severe reading problems. If these students were to fail the third grade State Student Assessment Test, the law mandates that they be retained, unless the skills are remediated. If 1,000 students are retained, at an annual cost of over $2,000 per pupil (Wellborn, 1979), the cost of this program would be in excess of $2 million per year for that county. It is readily apparent that a program which could facilitate the fluent learning of and the retention of skills could have considerable impact on public education.

---

*Florida Primary Education Act of 1979, CSHB 1036, Section 230.231.

Vergason (1964, 1968) wrote that retarded people typically have memory problems. He found that some differences in memory between retarded and normal people can be eliminated by providing extra practice or "overlearning" on specific skills for the retarded (Vergason, 1964). Procedures that help improve memory and retention in handicapped students would be a great help to the field of special education.

The research literature on learning and retention is large, and consistently supports the premise that degree of original learning is perhaps the single most important factor affecting retention. Early investigations of memory by Ebbinghaus (1885), Luh (1922), Krueger (1929, 1930), and Underwood (1954) laid the groundwork for much of the later work in this area. Results of this study will be compared and related to many of these earlier works.

There are several major problem areas and questions that have not been well-addressed. One of these is the specific nature of the relationship between degree of learning and retention. A "diminishing returns" effect for higher degrees of learning has occasionally been reported (Gilbert, 1957; Krueger, 1930; Luh, 1922; Postman, 1962). The problem of determining how much learning or how much fluency is necessary has not been solved.

An equally important question is to determine precisely the requirements for increasing degree of learning or fluency. Attempts to increase learning and retention have involved various measures. In some cases, extra time for study of a specific subject or topic has been provided. A more precise method has provided additional practice trials

on a specific skill. A third method has been specifically to increase performance scores. This pinpoints another area of concern; that is, how to best quantify performance. Are accuracy scores sufficient or are rate-of-response (frequency) measures necessary?

The basic issue involved is whether improved retention is due merely to providing extra time, or to extra practice, or if an actual increase in performance is necessary. There is a growing field in education called Precision Teaching, which primarily uses rate-of-response, or frequency, as a measurement unit. Various performance aims, in terms of frequency, have been suggested. Educators supporting the use of Precision Teaching generally agree that a high rate of response is necessary for the learning and retention of skills. An alternative view, while supporting the basic notion of direct instruction, suggests that practice alone is the major factor for retention. Neither the Precision Teaching literature nor the research literature concerning degree of learning has thoroughly investigated this question. In many cases, the distinction between performance change and additional practice has not even been made. This study will make that distinction and will attempt to separate the effects on retention due to extra practice, and those due to improved performance.

Statement of Problem

This study is designed to investigate the relation between frequency of response and retention of learned skills. Several questions were addressed. First, what is the nature of the relation between frequency of response and retention? For example, does frequency of

response on the performance measure affect the rate of forgetting? Second, does a program that specifically increases performance affect retention differently than a program that merely provides extra practice?

Limitations

A possible limitation to this study concerns the selection of subjects. There was no random selection or assignment of subjects or classrooms. Teachers in one school were asked to participate, and two of the teachers responding taught at the same grade level. While this may limit the generality of results concerning subjects, the choice of experimental task should help minimize this limitation.

Delimitations

In one respect, the experimental task used in this study may be viewed as rather different from most school tasks. It is this characteristic which helps control for interference from other school activities, since the content of the task (a trigram paired with a number) is rather unique. However, the stimulus-response combination used (visual-written) is common to many school tasks, and should help provide a basis for generalization to other skill areas.

A second possible delimitation is the length of the retention interval. Retention was measured after two-week periods. A longer interval may not yield similar results.

CHAPTER II

REVIEW OF RELATED LITERATURE

Scope of Review

There is a voluminous research literature concerning learning and retention. However, only those studies relative to the particular questions of this experiment were reviewed. Several criteria for selection were used. First, retention of learned material or learned behavior must be a major dependent variable. Second, degree of learning or practice must be a major independent variable. Third, only those studies using a laboratory or controlled type of setting were reviewed.

This last requirement was made to limit studies done in "applied" or school type settings, where experimental control is often not adequate. The daily routine and requirements of classroom instruction necessarily limit the degree of control that can be exercised in an experiment in that setting. Because of the confounding factors of teacher variability, subject matter differences, and poor variable definition and measurement, the applied literature was not extensively reviewed. However, several pertinent studies conducted in school settings have been included, because of the design and/or relevant data collected.

The use of "overlearning" procedures to enhance retention is a widely recommended practice, as Kratochwill, Demuth, and Conzemius (1977) point out. Yet, the basis for these many suggestions seems somewhat narrow. Relatively few studies have been consistently cited as supporting extra learning (usually, those studies with the term "overlearning" in the title). It is apparent that a comprehensive review of studies which have systematically varied degree of learning (or practice, or performance) has not been done.

By including only laboratory studies which evidenced acceptable experimental control, and by enlarging the review past the "overlearning" studies, to include studies which met the criteria concerning dependent and independent variables, a comprehensive body of research literature was assembled that relates specifically to the effects of degree of learning. From this literature, conclusions can be drawn concerning common procedures, results, and particular strengths and weaknesses. If consistent support for "overlearning" cannot be found in this literature, the support probably does not exist. However, if the literature reviewed does support the notion of "overlearning," then follow-up research (both applied and experimental) and instructional applications may be suggested.

Sources Used for Review

Initial references on overlearning are not difficult to find. The verbal learning studies by Krueger (1929), Gilbert (1957), Underwood and Keppel (1963), Vergason (1964) and Lance (1965) are often cited. Motor

learning studies which appear frequently as overlearning references include those by Krueger (1930), Melnick (1971) and Chasey (1971, 1977). References for the majority of the 37 studies reviewed here were obtained from searching reference lists of other articles. A search of the ERIC data base was conducted, using "degree of learning" and "overlearning" as descriptors. The Current Index to Journals in Education was also searched using the same descriptors.

Chronological Overview

Both verbal learning and motor learning tasks have been used in the study of retention. Between 1879-1884, Hermann Ebbinghaus conducted a series of experiments investigating memory and retention of verbal material. From these studies came a publication entitled Memory: A contribution to experimental psychology (Ebbinghaus, 1885).

This pioneering work laid the groundwork for a field of study that continues today. The Ebbinghaus studies are notable for a number of reasons. First, Ebbinghaus emphasized "The Method of Natural Science," insisting upon exact measurements. Second, Ebbinghaus invented the nonsense syllable as a unit of material. This provided a large number of homogeneous items that would not be affected by factors influencing the learning and retention of meaningful material. Third, Ebbinghaus developed the "savings score," which counted the number of trials or amount of time saved in relearning a list. This provided a retention measure that could be used where actual recall was very low or zero. Yet, the effects of the original learning would appear in relearning scores.

The precision with which Ebbinghaus carried out his work, particularly in relation to the general knowledge of learning available at the time, is remarkable. Features of this work continue to appear in and affect present day research. In the the 1920's and 1930's, degree of learning, or what was called "over-learning" (Luh, 1922), was used as the major independent variable in a progression of studies dealing with retention. Luh (1922) investigated the effects of degree of learning, retention interval, and method of recall on retention of verbal material. Krueger studied the effects of degree of learning and retention interval on retention of verbal material (Krueger, 1929), and on retention of a motor task (Krueger, 1930). McGeoch (1929) studied the effects of extra practice on inhibition.

During the 1940's, Grace Rubin-Rabson published a series of retention studies using piano-playing as the task. Several of the studies systematically varied the amount of practice given (Rubin-Rabson, 1941a, 1941b). Also during this period, a pursuit-rotor task was used to study effects of amount of practice (Buxton, 1943; Irion, 1949; Jahnke, 1958).

In the 1950's, verbal tasks were again used in retention studies. Gilbert (1957) and Slamecka (1959) measured recall of facts from a prose selection. Underwood (1954) and Underwood and Keppel (1963) used a paired-associates task, while others (Rock, 1958; Postman, 1962) used syllables as the retention task in overlearning studies. Peterson and Peterson (1959) and Hellyer (1962) used a three letter trigram and retention intervals of seconds.

During the 1960's and 1970's several new directions were taken in studies varying degree of original learning. Ammons, Farr, Bloch,

Neumann, Dey, Marion, and Ammons (1958), Fleishman and Parker (1962), and Hammerton (1963) studied retention of complex perceptual-motor tasks. Vergason (1964) and Lance (1965) used a paired-associate task in comparing memory of normal and retarded subjects, while varying degree of original learning. Chasey (1971) used a diagonal hop to study retention of motor skills by retarded subjects. Melnick (1971) studied performance on a stabilometer by college students, and Chasey (1977) used a stabilometer with retarded boys. Chasey and Knowles (1973) measured retention of a bean bag toss with retarded boys. Roberts (1972) studied performance of pigeons on a delayed matching to sample task, a procedure analogous to retention studies with humans. Brehmer and Lindberg (1973) varied degree of learning in a single-cue probability task. Kratochwill et al. (1977) studied the learning of sight words by pre-school children.

Subjects, Intervals, and Related Issues

Studies investigating the effects of degree of learning have varied widely regarding task, retention interval, subjects, and degree of learning. Some researchers have used college students or faculty as subjects (Ammons et al., 1958; Brehmer & Lindberg, 1973; Buxton, 1943; Fleishman & Parker, 1962; Hammerton, 1963; Irion, 1949; Krueger, 1929, 1930; Luh, 1922; Melnick, 1971; Peterson & Peterson, 1959; Underwood, 1954; Underwood & Keppel, 1963). Ebbinghaus (1885) used himself as a subject, while Rubin-Rabson (1941a, 1941b) used piano teachers, Gilbert (1957) used soldiers, Hellyer (1962) used housewives, and Roberts (1972)

used pigeons. Chasey (1971, 1977), Chasey and Knowles (1973), Lance (1965) and Vergason (1964) used retarded subjects, while Kratochwill et al. (1977) used pre-school children.

Retention intervals have varied from a period of seconds (Hellyer, 1962; Peterson & Peterson, 1959), to minutes (Buxton, 1943; Irion, 1949; Jahnke, 1958; McGeoch, 1929; Slamecka, 1959), to days (Gilbert, 1957; Krueger, 1929, 1930; Luh, 1922), to weeks and months (Ammons et al., 1958; Chasey, 1971, 1977; Chasey & Knowles, 1973; Fleishman & Parker, 1962; Hammerton, 1963; Melnick, 1971; Rubin-Rabson, 1941a, 1941b). In some cases, retention intervals exceeded one year in length (Ammons et al., 1958; Fleishman & Parker, 1962).

There are a number of factors besides degree of learning that affect retention. These factors, including meaningfulness of material, type of practice, method of recall, similarity of items, mediation, and interference, have been discussed at length by a number of authors (Hall, 1966; Hovland, 1951; McGeoch and Irion, 1965; Underwood, 1964; Vergason, 1968). Underwood and Keppel (1963) point out the need to control for level of association or meaningfulness when using verbal tasks. Gillette (1936) suggested using an "adjusted learning" procedure to control for intra-list differences. In this procedure, an item or pair is dropped from the learning trials when it has been recalled once correctly. This prevents a differential number of trials beyond initial recall from being given for items within a list. However, degree of learning has consistently been shown to be strongly related to retention. Reviews of retention studies have emphasized the importance of

degree of learning (Hilgard, 1951; Hovland, 1951; Underwood, 1964a, 1964b; Vergason, 1968). Since degree of learning is directly related to retention, it has been used, both explicitly and implicitly, as the dependent measure in learning and retention studies. In discussing variables affecting retention, Underwood (1964b) wrote "If we are to assess properly the influence of these variables on retention, degree of learning must be equated, since degree of learning and retention are directly related" (p. 128).

Quantification of Degree of Learning

In all of the studies in this review, degree of learning or amount of practice provided was a major independent variable. Several procedures have been used to quantify this variable. One commonly used method is to calculate average number of trials needed to reach 100% accuracy on the task. This number is 100% learning, or 0% overlearning, and is the basis for other levels used. It is increased or decreased to produce various degrees of learning. Dividing the number of trials in half, then, gives 50% learning. Multiplying the base figure by 2 yields 200% learning, or 100% overlearning. This procedure was originated by Luh (1922), and has been commonly used in verbal learning studies (Gilbert, 1957; Krueger, 1929; Postman, 1962; Underwood, 1954). It has also been used in studies of motor learning (Brehmer & Lindberg, 1973; Chasey, 1971, 1977; Krueger, 1930; Melnick, 1971; Rubin-Rabson, 1941a, 1941b).

Some researchers used number of trials or repetitions as a measure with no reference to a "100% learning" level. (Ammons et al., 1958;

Ebbinghaus, 1885; Hellyer, 1962; Irion, 1949; Roberts, 1972; Underwood & Keppel, 1963). The "overlearning" and "learning" designations can be confusing, but both have been used, sometimes in the same study (Krueger, 1929). An alternative to the "number of trials" procedure is the setting of specific performance levels as criteria. Buxton (1943) used seconds on target as a criterion score. Fleishman & Parker (1962) and Hammerton (1963) also used performance-based measures as criteria.

At one level of analysis, there may seem to be little difference between the performance based measures and those based on 100% learning. Increases up to 100% learning in this latter type of score will most likely result in increased performance, based on accuracy of response. However, if a paired-associates list were learned perfectly with only half the base number of trials, there is no measure of improved performance available. Likewise, when trials are given beyond 100% learning, increased performance can only be assumed.

The procedure used by Buxton (1943), however, used a direct measure of performance. He measured time on target for a pursuit-rotor task. Criterion scores of 90 seconds, 360 seconds, and 630 seconds were used. Since these are direct performance measures, no assumptions concerning improved performance are necessary.

Type of score used is closely related to the issue of criteria. In procedures where the only measure of performance is accuracy of response, performance levels are necessarily limited to a top of 100% correct. This "ceiling" will affect scores for the so-called "overlearning" levels. When direct performance measures such as frequency of response or time-on-target are used, the problem of imposing a ceiling effect can be avoided.

Two scores commonly used in the studies of verbal learning are number or percent of items recalled correctly, and a "savings score" (Ebbinghaus, 1885). In the savings method, number of trials or time required to relearn are subtracted from number of trials or time needed for original learning. This difference score is then divided by the score for original learning, yielding a percent of savings. The formula for calculating a savings score is:

$$\frac{OL-RL}{OL} = \text{savings}$$

Several authors have discussed the measurement problems related to learning and retention studies. Hilgard (1934) illustrated discrepancies in savings scores caused by counting or not counting the "test trials" in learning or relearning trials. He favored exclusion of test trials, while McGeoch and Irion (1965) favored including the learning test trial, but excluding the relearning test trial.

Underwood (1964) pointed out that if number of trials is used as the criterion, amount of learning by individuals and for different materials are not necessarily equated. He suggested use of the "single-entry projection" technique. In this procedure, a group learning curve is constructed, and a hypothetical score is projected for use as a test trial. This eliminates the problem of counting or not counting test trials. A similar procedure, using pooled data, is proposed for scores based on criterion performance. This procedure does not eliminate the problems inherent in using derived scores, however (McGeoch & Irion, 1965). In one study using the single-entry projection technique, the

authors blamed the "fortunes of randomization" for the fact that their calculations produced a number of recall scores in excess of 100% (Underwood & Keppel, 1963).

There seem to be several other possible solutions to the measurement problem. The first is related to experimental design. Some researchers exposed all groups to each experimental condition (Luh, 1922; Roberts, 1972), while others used separate groups for each condition (cf. Krueger, 1929, 1930; Vergason, 1964). In all studies except that of Ebbinghaus (1885), group data are analyzed based on average scores. Averaged data cannot account for within group differences (Johnston & Pennypacker, 1980; Sidman, 1960). Despite this major disadvantage of using group data, results from the overlearning studies are quite consistent. Only Rubin-Rabson (1941a, 1941b), Brehmer and Lindberg (1973), and Reynolds and Glaser (1964) found no differences due to extra practice.

A second possible solution is the avoidance of savings scores as a measure of retention. The problems associated with the various ways of calculating this score have been pointed out. The use of projected scores based on group data seems hardly an appropriate solution. Some of the studies reviewed used such indirect or derived scores for retention measures (Krueger, 1929, 1930; Underwood & Keppel, 1963). Perhaps the best solution is to rely on the most direct measures of performance available.

Hilgard (1951) discussed eight points that are crucial in retention studies. Two of these items concern the scores to be recorded and the criterion for mastery. Certainly the choice of measures is vital to the

validity of the investigation. The problem of mastery criterion, however, can be avoided by choosing a measure that has no ceiling effect. The designation of mastery level is subjective, and irrelevant in this context. Little evidence was found during this review that Skinner's concept of frequency of response (rate) was used as a measure of performance. Skinner (1953) has identified seven advantages of rate data. First, rate data are orderly, they follow the laws of behavior. Second, rate data refer to individual subjects. Third, rate studies are easy to reproduce. Fourth, rate data provide continuous measurement. Fifth, this system lends itself to experimentation. Sixth, it provides a basis for probability. Seventh, frequency of response is a very sensitive measure.

Some of the studies reviewed used a time factor in measuring learning. Ebbinghaus (1885) recorded total elapsed time to learn a list. Luh (1922) also recorded time, but most data presented were in terms of percentage of items recalled. Studies of motor learning have often used direct, time-related measures. In pursuit-rotor tasks (Buxton, 1943; Irion, 1949; Jahnke, 1958) time on target, was recorded. Stabilometer studies (Chasey, 1977; Melnick, 1971) used time in balance as a measure. Several studies using a complex perceptual-motor task have used time related measures, and based criterion scores on performance, rather than on number of trials (Fleishman & Parker, 1962; Hammerton, 1963).

The difference between improved performance and extra practice is one of the major points of this review. A key factor is whether improved retention is due merely to extra practice, or if an actual change

in performance is necessary. This issue has not been well-addressed in the literature. In fact, most authors have not distinguished practice from performance. Since the study by Luh (1922) when the term "overlearning" was introduced, many researchers have assumed that merely providing extra practice would, in fact, increase performance. In most cases, this is strictly an assumption. If retention increased due to extra practice, as it did in most studies, the conclusion drawn was that overlearning was effective. However, in the studies reviewed which showed no differences (Brehmer & Lindberg, 1973; Reynolds & Glaser, 1964; Rubin-Rabson, 1941a, 1941b) the conclusion was that overlearning was not effective. The appropriate conclusion may have been that the extra practice did not improve performance, nor retention. Therefore, extra practice was ineffective.

There are indications, however, that extra practice can be effective, even when no increase in performance could be directly measured. Many of the "overlearning studies" are in this category (Chasey, 1971, 1977; Chasey & Knowles, 1973; Gilbert, 1957; Knowles, 1973; Krueger, 1929, 1930; Lance, 1965; Melnick, 1971; Postman, 1962; Underwood, 1954; Vergason, 1964). Some of the authors of motor learning studies have described the independent variable in terms of trials or repetitions, and not used the terms learning or overlearning at all (Ammons et al., 1958; Buxton, 1943; Irion, 1949; Jahnke, 1958; Roberts, 1972). This would seem to be a more precise procedure.

Individual Studies

In reviewing these studies, an attempt was made to determine if an actual change in performance occurred, or if extra practice was the only difference. It has been noted that very few investigators used a performance measure as a criterion score. However, in some cases a change in performance could be noted from the data, even though the authors may have used practice trials as criterion scores.

The studies reviewed here, all of which used degree of learning or amount of practice as an independent variable, overwhelmingly support the conclusion that retention increases with practice or performance. The studies which showed no beneficial effects of extra practice were analyzed in an attempt to determine the reason for such results.

Studies were analyzed relative to three major questions. First, what were the author's conclusions, and are they supported by the data? Second, are results due merely to increased practice, or was a change in performance also evident? And third, what is the specific nature of the relation between retention and practice/performance? In the summary tables (Appendix L), the results section indicates whether there was an actual performance change, and the nature or the retention curves. In most cases, this analysis required re-evaluation of the author's data, including charting of the data on semi-log graph paper. The AC-4 chart (Berquam, 1979) was used for this purpose (See Appendix A). The advantages of this type of chart or graph are discussed at length in several publications (Johnston & Pennypacker, 1980; Pennypacker, Koenig, & Lindsley, 1972; Schmid, 1954). One characteristic of such data

presentation is that a straight line indicates a constant rate of percentage of change. Also, sets of parallel lines indicate equivalent percentage or proportional differences at any point along the lines.

The chart in Appendix B shows three basic relationships. Picture 1 shows parallel lines. As A changes, B changes in the same proportion. Picture 2 indicates that as C changes, D changes in a manner that produces an increase in the proportional differences between C and D. Picture 3 indicates that as E changes, F changes in a manner that produces a decrease in the proportional differences between E and F.

With retention curves, parallel lines would indicate that retention is a constant percentage of previous performance or practice level. Divergent lines, as in Picture 2, would indicate that greater performance is related to a greater percentage of material retained. Picture 3, showing convergent lines, would indicate that greater performance is related to a smaller percentage of material retained. For the purposes of this paper, divergent lines will be considered to show proportionally greater or proportionally increasing retention due to practice, while parallel lines will be classified as proportionally equivalent retention due to practice. Divergent lines will be described as showing proportionally decreasing retention due to practice.

Verbal learning studies

In the inaugural study of retention, Ebbinghaus (1885), using himself as subject, learned lists of non-sense syllables. In a series of investigations extending over several years, Ebbinghaus systematically varied length of list, number of repetitions in learning,

retention interval, number of relearnings, time of day for learning, and order of items in each list. Time and number of trials to criterion were used as dependent measures. The criterion used was one complete recall. Retention, then, was measured as a savings score, or percent of time or trials saved in relearning a list. Number of extra repetitions was increased, providing 8, 16, 24, 32, 42, 53, and 64 repetitions.

Seconds saved in relearning the lists increased in proportion with number of repetitions. When repetitions needed to relearn a list on successive days were counted, there were substantial decreases each day. Percent of savings during relearning was greater for longer lists.

C. W. Luh, working at the University of Chicago, published a lengthy monograph entitled The Conditions of Retention, (Luh, 1922). This report included information concerning the effects on retention due to degree of learning, method of recall, and retention interval. Degree of learning was based on average number of trials to learn a list to a criterion of one correct recall. This was termed 100% learning. Other degrees of learning were based on this figure.

This study supported the conclusions of Ebbinghaus (1885), that retention decreases as a function of time interval. The basic negatively accelerated curve as reported by Ebbinghaus was obtained. Luh's subjects, ten college students and professors, learned separate lists of syllables to four degrees of learning. Retention over seven intervals ranging from 20 minutes to 48 hours was measured. Luh reported that as degree of learning increased, retention increased. He also stated that at the higher degrees of learning, the proportional increase in retention as measured by savings scores, was small. This was the first

mention of a "diminishing returns" effect in retention studies. However, when absolute retention scores are graphed on the AC-4, there is less proportional decline for the higher degrees of learning, up to 100% (See Appendix F). This analysis indicates, in contrast to Luh's conclusion, that higher degrees of learning, up to the 100% level, are more effective in promoting retention. The comparison of savings scores can hide this effect. Luh based his degrees of learning on number of practice trials given, a common procedure in overlearning studies. Individual practice trials can have differential effects on performance or learning, based on the degree or stage of the learning (Ebbinghaus, 1885; Haring et al., 1978). This may also partially account for Luh's finding of a diminishing returns effect.

Krueger (1929) attempted to replicate and extend the work of Luh. He studied the effects of three degrees of learning on retention over intervals of from one to 28 days. Krueger concluded that higher degrees of learning promote greater retention. This is in agreement with the re-analysis of the Luh data.

Using lists of non-sense syllables, McGeoch (1929) studied the effects of number of repetitions on retroactive inhibition. The study provided an opportunity to examine effects on retention due to degree of learning. Ten college faculty and students were given 6, 11, 16, 21, or 26 presentations on lists of non-sense syllables. Retention was checked after five minutes, using number of items correctly recalled and number of relearning trials. The experimental group learned a separate list between the learning and relearning tasks for the original list. In

this way, the effects of degree of learning on retroactive inhibition could be assessed. For both groups, recall and relearning scores increased with more repetitions. Amount of inhibition, as measured by recall and relearning scores, decreased as repetitions increased.

Underwood (1954) investigated the effects of three degrees of learning, while also varying material (serial learning and paired-associate tasks), and practice format (distributed and massed practice). Degrees of learning were 50% and 100% learning (based on accuracy) and 700% overlearning (seven additional trials after perfect recall). Thus, there is a direct measure of performance for the 50% and 100% learning groups. Comparisons on the AC-4 of average number of items recalled after 24 hours for each group indicate that retention was proportionally greater as degree of learning increased (See Appendix H). The author concluded that 1) no interaction was found between degree of learning and spacing of practice; 2) no differences were found due to spacing of practice; 3) for relearning, significant differences were found only for degree of learning.

Several studies dealing with the effects of extra practice on retroactive inhibition have been done. Richardson (1956) used groups of "concept" words for the learning task. He found inhibition to decrease with an increase in practice trials, but found no improvement in retention due only to the extra practice. Analysis of the retention scores reported indicates that they were all very near the maximum possible score. Thus, little improvement was possible. In another study involving inhibition, Briggs (1957) found performance and retention both increased with additional practice, while amount of inhibition decreased.

Slamecka (1959) studied the effects of repeated practice on the retention of connected discourse. Using 140 college students as subjects, he gave 1, 3, 5, 6, 7, or 9 presentations of a 28 word sentence, narrated at a rate of 60 words per minute. Original learning (performance) and retention scores, based on words recalled correctly, increased with more repetitions. Retention after 5 minutes was proportionally greater with more repetitions.

Gilbert (1957) used recall of facts from a short story read to the subjects as a retention task. Two degrees of learning over retention intervals of 15 minutes, 24 and 48 hours were studied. Main effects for interval and degree of learning were found, with no interactions. Retention curves for degree of learning over time were proportional, indicating that retention was proportionally equivalent across degrees of learning.

Peterson and Peterson (1959) and Hellyer (1962) used a rather unique task, with very short intervals, to assess the effects of practice. The task involved telling the subject a three-letter trigram. Then, after a period of several seconds for rehearsal, the subjects were required to count backwards for a specified time. At the end of this period, the subject attempted to recall the original item. The authors concluded that recall decreased as retention interval increased, and that recall increased with rehearsal time.

Hellyer (1962) used four levels of practice or rehearsal, and four retention intervals. Results were similar to those reported by Peterson and Peterson (1959). In addition, a significant interaction between the two variables was found. When the author's data are graphed on the

AC-4, the interaction is apparent. As delay interval increased, additional repetitions produced proportionally greater recall (See Appendix I).

A number of investigators have studied the role of repetition in associative learning (Estes et al., 1960; Rock, 1958; Wollen, 1962). Using a verbal learning task involving paired-associates, the major question in these studies was whether associations (learning) of individual items occur gradually over trials, or all at once on a specific trial. There are differences in the conclusions drawn concerning the basic question of incremental or one-trial learning. The Estes et al. (1960) conclusions support one trial learning, as do those of Rock (1958). Wollen (1962) concluded that learning occurs incrementally. Each of the studies provided an opportunity to assess the effects of repetition on retention. Conclusions regarding this relationship were consistent. Retention scores increased directly with number of trials given for learning (Estes et al., 1960; Rock, 1958; Wollen, 1962).

In a very brief description of an overlearning study, Postman (1962) investigated the effects of three degrees of learning over one retention interval. He used lists of high-frequency and low-frequency words. Overlearning levels of 0%, 50%, and 100%, were used. Highest retention scores were reported for 100% overlearning, with little difference found between 0% and 50% overlearning. Differences in absolute retention scores were small. Differences in relearning scores were significant at each level.

Underwood and Keppel (1963) varied degree of learning, association level of material, and retention interval. Letters were used in a

paired-associates paradigm. The authors concluded that the only variables which substantially affected retention are interval and degree of learning. They found no differences due to association level of material, when degree of learning was equivalent. When data from the author's Table 1 (mean correct responses at each condition) are graphed on the AC-4, several relationships are apparent. First, retention decreased over time. Second, retention changed little as a function of association level when degrees of learning were equivalent. Third, retention over time was generally proportionally equivalent for degree of learning, although there were some indications of proportionally increasing retention (See Appendix J).

Reynolds and Glaser (1964), in a study of the effects of extra repetitions in the learning of programmed science material, found few differences in retention after two days or after three weeks. However, most scores were near the ceiling level; thus, little room for growth was available.

Several investigators have studied retention and degree of learning with retarded subjects. Vergason (1964) compared retention scores of retarded and normal students, using two levels of learning, with retention intervals of one and 30 days. Pictures were used in a paired-associates task. Original learning scores showed significant differences only for level of learning. Vergason concluded that rate of learning for the normal and retarded subjects was equivalent. For the one-day recall scores, normal subjects were superior to retarded subjects

at both degrees of learning, and overlearning produced higher scores for both groups. At the 30 day interval, there was an interaction between intelligence and degree of learning. Vergason concluded that at lower levels of learning, retarded subjects had poorer retention. However, when material was overlearned, and the retention interval was longer, retarded subjects had scores which were equivalent to those of normal subjects.

Lance (1965) also compared retention of retarded subjects to that of normals. Two levels of meaningfulness and two levels of learning were studied. A paired-associates task was used. The only difference in retention as measured by a savings score was due to degree of learning. No differences were found due to meaningfulness or intelligence designation of subjects.

Kratochwill et al. (1977) taught a list of four common words to pre-school students. The criterion used for the control group was two consecutive, errorless trials. One experimental group received 10 extra trials, and a second experimental group received 20 extra trials. Recall was checked after 60 minutes. Analysis of variance yielded significant differences only for the 20 trial group, but a trend analysis indicated that retention did improve with extra practice. The small number of items learned (four), and the procedures used in this study (massed practice; a single, brief retention interval) may account for the limited effects found.

Unpublished data from a major precision teaching project in Great Falls, Montana (Sacajawea Project), reports frequency data from first through sixth graders, on several academic skills. Performance data

from spring and fall are presented (Clement, 1978). Analysis of 25 data sets indicates that 15 instances showed retention to be basically proportional to performance. Nine data sets yield results showing retention to be proportionally greater at higher frequencies. One data set showed proportional retention decreasing as a function of frequency (See Appendix K).

Motor learning studies

William Krueger, in one of the earliest overlearning studies involving a motor learning task (Krueger, 1930), used a finger tracing maze with a group of 192 college students. Retention intervals varied from one to 28 days. Degrees of learning used were 100% learning, and 50%, 100%, and 200% overlearning, based on number of trials needed to first correctly trace the maze. Data are presented concerning number of trials needed to learn and relearn a maze, the difference in time required to learn and relearn, and the difference in errors made in learning and relearning. Krueger concluded that higher degrees of learning always produced higher retention scores, although the change from 100% to 200% overlearning was not as "economical" as the change from 50% to 100% overlearning.

In the late 1930's and early 1940's, Grace Rubin-Rabson investigated the effects of several aspects of practice on piano-playing performance. Two of these studies used degree of learning as an independent variable (Rubin-Rabson, 1941a, 1941b). Using accomplished pianists as subjects, Rubin-Rabson provided various practice formats on short, individual pieces of music. Retention was measured after two

weeks and again after seven months by recording how many trials were required to relearn each passage to the criterion of one smooth recital.

In the first study, Rubin-Rabson (1941a) provided three types of rehearsal, one of which resulted in approximately 100% overlearning, based on number of trials to criterion. The other two procedures provided mental rehearsal; in one condition about midway in the learning period, and in the other, after criterion had been reached. Relearning scores at both intervals showed no significant differences due to method of practice.

In a second study investigating overlearning, Rubin-Rabson (1941b) used three degrees of overlearning, again based on number of trials to criterion. Retention intervals were two weeks and seven months. Clear conclusions from this study are difficult to draw, because of several complications. However, relearning scores did not vary significantly due to amount of learning.

Several weaknesses of the Rubin-Rabson studies need to be pointed out. First, the nine passages learned by each of the nine subjects varied in difficulty. Also, the degree of skill of each of the musician subjects varied. The rather primitive analysis of variance used indicated significant interaction for materials and for subjects. This study also shows one of the problems in using number of trials and a high criterion as the basis for overlearning measures. Additional practice can be provided, but practice has less effect on strong performance than on weak performance (Ebbinghaus, 1885).

Another motor task that has been used in overlearning studies is tracking a target on a pursuit rotor (Buxton, 1943; Irion, 1949; Jahnke,

1958). In this task, a brass target spins on a disc, and the subject attempts to keep a stylus in contact with the moving target. Contact with the target actuates a timer, so a direct measure of performance is available. Retention intervals are typically quite short, with a reminiscence effect often found. Buxton (1943), using college students as subjects, found performance scores increased directly with degree of learning. He also found reminiscence to occur after a ten minute rest. Greatest proportional reminiscence occurred for the lowest degree of learning and decreased as additional practice was provided.

Irion (1949) also used college students in a pursuit-rotor task involving overlearning. He found performance to be directly related to amount of practice. Also, proportional reminiscence was greatest for the lowest performance level, and least for the highest performance level.

Jahnke (1958) studied motor learning on the pursuit rotor, using four levels of learning and three rest intervals, with high school students as subjects. Results of this study are in agreement with earlier reports (Buxton, 1943; Irion, 1949), showing performance directly related to amount of practice, and proportional reminiscence inversely related to performance. Also, retention curves over the one day and seven day intervals were proportional for amount of practice.

Several investigators have studied the effects of degree of learning on the performance and retention of complex motor skills. Ammons et al. (1958) used a task requiring sequential manipulation of a variety of switches and knobs. Fleishman and Parker (1962) and Hammerton (1963) used tasks simulating a radar tracking procedure, where subjects were

required to move a dot on a cathode ray tube to a target area. Ammons et al. (1958) studied the effects of two degrees of learning of a sequential motor task, over six retention intervals ranging from one minute to two years. University students served as subjects. Individual groups were used for each experimental condition, and were matched by beginning performance. Performance was found to decrease over time, and increase with degree of learning. Trials required to relearn the task decreased with degree of learning and increased with retention interval. Retention was proportionally greater over time for the higher performance group. No interactions were found.

Fleishman and Parker (1962), in a lengthy and complex study designed to investigate various practice effects on retention of a complicated perceptual-motor task, trained subjects in a task simulating a radar intercept mission for a pilot. Major independent variables included type of training, and retention interval. Retention intervals ranged from one to twenty-four months.

Little data or information are presented concerning initial proficiency levels. However, there apparently were enough differences shown for the authors to conclude that level of initial proficiency is the most important factor influencing retention. The authors present correlation coefficients between original proficiency and retention. They state that there was no change in the ordering of groups over time.

Hammerton (1963) used a tracking task similar to that used by Fleishman and Parker (1962). The effects of two degrees of learning over a retention interval of six months were assessed. Eighteen adult university professionals served as subjects. No difference was found

between the control and experimental (overlearning) groups on initial learning. Significant differences were found between groups on the first retention trial, but relearning measures were quite similar. The author concluded that if performance on the initial trial after a no-practice interval is important, overlearning is beneficial. However, if time is available for a period of warm-up practice after the interval, then overlearning may not be necessary.

A major strength of the above studies was the measure used for quantifying performance and retention. Because of the nature of the tasks used, direct, time-related measures were available; thus, there was no need to resort to indirect, derived, or projected scores.

Purdy and Lockhart (1962) investigated the long term retention of novel motor skills by 36 college women. Performance at the end of a learning period on five toss-catch and balancing tasks was designated as high, average, or low. The specific type of scores used was not presented. Retention and relearning scores were compared after approximately one year. All scores were converted to standard scores. The authors concluded that the average group retained a greater proportion of their original learning scores than did either the high or low groups. Actual differences in percentage retained were small. When median percentage retained for each group and by each skill is compared, there is a direct relationship between skill level and retention, thus indicating that averages were probably strongly influenced by extreme scores.

In investigations of the effects of practice on motor learning, several researchers have used a balancing task. The task used in these

studies required subjects to balance on a device called a stabilometer. This is a board placed on a center roller. Electronic timing devices measure time-in-balance during each trial, thus providing a direct measure of performance.

Melnick (1971) used college students for subjects, and found that each of three overlearning degrees produced recall scores that were superior to scores for the control group. However, relearning scores were quite similar for all groups. Only the 200% overlearning group had a significantly superior relearning score. Recall and relearning scores after eight weeks were found to be directly related to degree of learning on the original task. Chasey (1971) reported results supporting an overlearning procedure in another study with the retarded. In this task, the subjects were required to hop diagonally to a specified square on a grid. When subjects were required to relearn the task after four weeks of no practice, the overlearning group had superior results.

Each of the three studies cited (Chasey, 1971, 1977; Melnick, 1971) used separate groups for each experimental condition. However, the groups were compared on original learning scores in each of the studies, with no significant differences found. Ceiling effects due to the use of practice trials for criterion scores rather than performance measures may account for the limited effects reported by Melnick (1971).

Two other studies relating to overlearning do not fit into either the verbal learning or motor learning categories. Roberts (1972) used a delayed matching to sample task with pigeons. This is similar to a retention study with humans. By requiring 1, 5, or 15 repetitions on

the sample before switching to the match, "overlearning" procedures were incorporated. Results were similar to many human retention study results. Retention decreased as the delay period increased. Increasing repetitions increased retention, and retention was proportionally equivalent as repetitions were increased. Brehmer and Lindberg (1973) used a single-cue probability task, where the subjects attempt to predict an item, based on previous items. The "validity" of previous cues was varied, as was amount of practice provided. Learning levels of 100%, 200%, and 300% were used, and no differences were found due to degree of learning. This task is similar to what might be called a "concept formation" task. Once the concept is learned, little difference in performance will be apparent on individual items. This factor imparts a ceiling effect that limits performance increases.

Summary

In summarizing results of the studies reviewed, a summary table format was developed (See Appendix L). Studies were first classified as verbal or motor learning studies. Then, individual studies were listed and described in terms of setting variables. Number and type of subjects, type of task, and length of retention interval were described. In the column headed "Degrees," the amount of learning or practice provided was listed. An attempt was made to make the descriptions as consistent as possible. Therefore, a specific description may be somewhat different from what the original author had used.

In the "Results" section, conclusions listed are those drawn from a re-analysis of the original published data, when possible. Three specific questions were addressed for this review. First, do results support overlearning procedures as improving retention. Second, was a measurable change in performance evident as a result of the manipulation of the major independent variable? Third, can a conclusion regarding the form of the relation between learning and retention be drawn?

In all, 37 studies were reviewed. Twenty-nine found increased retention for the experimental group, and four studies had mixed results. Eleven studies provided some measure of performance change in learning scores. All 11 of these studies reported increased retention for the higher-performing groups. In comparing retention curves between degrees of learning, data from 7 studies indicated proportionally greater retention for higher degrees of learning. Seven studies reported data that showed proportionally equivalent retention. Two studies indicated mixed results, with both proportionally greater and proportionally equivalent curves. No studies indicated that the effect of degree of learning on retention decreased over time, although differences in number of trials to relearn have decreased over time (Krueger, 1930), and one of 25 data sets in another study (Clement, 1978) showed this result.

The phenomenon of "diminishing returns" affecting higher degrees of overlearning has been discussed by several authors (Gilbert, 1957; Krueger, 1930; Luh, 1922; Postman, 1962). There appears to be con-

siderable confusion concerning what constitutes a diminishing returns effect. Luh (1922) reported actual decreases in number of items recalled for 150% learning. Krueger (1929, 1930) reported higher degrees of learning to be less economical, based on the fact that the ratio of number of items recalled at 200% learning was less than the ratio of practice trials given (2:1). This is a different result than that obtained by Luh (1922), and Krueger did not use the term diminishing returns. However, Gilbert (1957), McGeoch and Irion (1965), and Postman (1962) described the Krueger studies as evidencing diminishing returns.

It is clear that a consistent definition of diminishing returns in the overlearning studies has not been made. It is also clear that retention curves have consistently shown to be either proportionally equivalent or proportionally greater over time, as learning or practice was increased, when charted as in the re-analysis done here.

In the four studies reporting no advantage to overlearning (Brehmer and Lindberg, 1973; Reynolds and Glaser, 1964; Rubin-Rabson, 1941a, 1941b), performance on the tasks used was apparently close enough to maximum at the lower degrees that little or no increase in performance was possible from the extra practice provided.

To summarize the literature on the effects of degree of learning on retention, the following major points can be made:

1. Consistent definitions, scores, and measurement procedures have not been used.
2. Extra practice was nearly always beneficial; improved performance was always beneficial.

3. The effects of degree of learning on retention did not decrease, and often increased over time.
4. Degree of learning has been shown to be such a powerful variable that Underwood wrote "In short, when it comes to forgetting or remembering, what counts is not the nature of the material but the degree of original learning." (Underwood, 1964a, p. 92).

CHAPTER III

METHOD

This study is designed to assess the relation between frequency of response and retention of learned material. A paired-associates learning task was used. Several specific questions concerning practice procedures and performance measures were addressed.

Several terms describing the experimental and control conditions need to be introduced and defined. Under the experimental condition, subjects were given practice that specifically promotes frequency of response improvement. This was termed "practiced fluency." Under the control condition, unpaced practice was provided, with no emphasis on frequency of response. However, all behavior can be measured in terms of frequency. At the end of the unpaced practice phase, a performance measure using frequency of response was taken. Therefore, a common measure of performance was available for both conditions. Since no specific effort was made to practice for or increase frequency under the control conditions, this measure will be termed "incidental fluency."

Experimental Questions

This study attempted to answer three specific questions. First, what is the specific nature of the relation between frequency of response and retention? Second, are performance, retention and celeration

scores produced by the "practiced fluency" condition different from those produced by the "incidental fluency" condition? Third, for a given frequency on the performance measure, are there differences between conditions on the retention and celeration measures?

Independent Variables

The major independent variables for this study are training method and retention interval. The two retention intervals were held constant, at a period of two weeks. Frequency of response was allowed to vary and was measured periodically using one-minute performance samples. In these samples, frequency of correct and incorrect responses was recorded.

Dependent Variable

Retention was measured using the same performance sampling technique as described above. Both correct and incorrect responses were recorded. Repeated measures were not taken during the retention interval, in an attempt to control for practice effect during this phase of the study.

Units of Measurement

All performance samples were measured as correct and incorrect responses per minute. This frequency of response measure provides a sensitive and scientific measure of performance (Skinner, 1953). Advantages of this type of measurement have been discussed previously.

A measure of change in frequency was also used. This measure, termed celeration (Pennypacker, Koenig and Lindsley, 1972), is generally calculated as percentage of change in frequency per week. For this study, the change measure was percentage of frequency retained.

Subjects

For this study, two classes of third graders were used as subjects. The classes were chosen from a regular elementary school in Orange County, Florida. An attempt was made to select "typical" or "normal" classrooms. Descriptive data from achievement test scores were compared, to determine the degree of similarity between the two classes. These data will be presented in the Results section.

Experimental Task

A paired associates task was used for this study. A three-letter trigram was paired with a single digit number. The trigrams used were chosen from a list developed and used in a related study by Rock (1958). The task involved a visual presentation of the trigram, with the subjects making a written response of the appropriate number. This stimulus-response combination is similar to that commonly used in school settings, particularly in arithmetic. Yet, the content of the pairs is sufficiently unique that interference from other school activities is limited. Preliminary work (Berquam, 1981) showed that performance on this type of task is only slightly correlated with academic achievement (see Appendix N). This would indicate that previous academic success is not necessarily a prerequisite for successful performance on this task.

During the initial learning stage, cued practice was given. A correct sample of each pair was presented at the top of the practice sheet. During the accuracy and fluency practice sessions, no cues were provided. Samples of these sheets are in Appendix M.

Procedure

This study involved four phases: 1) Initial learning; 2) Retention I; 3) Relearning; 4) Retention II. The initial learning was divided into two stages. In the first stage, using cued practice, the students were required to do a matching to sample procedure. The students were given daily practice, using a practice sheet with the trigram-number pairs at the top. The instructions were to complete the page of 100 responses, writing the correct number for each printed trigram. Following this practice each day, an accuracy check was done, with the students required to write the numbers to each of the 10 trigrams without the cues being provided.

When the median percent of accuracy for the groups reached 100%, separate practice procedures were implemented for each of the two classes. The control condition provided daily practice in the same manner as in the previous procedure, but without the cues being given. The experimental condition provided daily one minute practice, with an emphasis on improving frequency of response. Each group was given five practice sessions during this stage. Thus, the number of practice sessions under control and experimental conditions were equivalent. However, under the control conditions, subjects completed the entire

practice sheet, for a total of 100 responses each day. Subjects in the experimental condition worked for only one minute each day. Since frequency measures during this stage ranged from approximately 20 correct to 80 correct responses per minute, the experimental condition provided fewer total responses than did the control condition.

At the end of this learning stage, each group was given a one minute performance check, and frequency of correct and incorrect responses was recorded. A two week retention period was then scheduled, with a one minute frequency check done at the end of the second week.

Both groups were then given five days of daily one minute practice sessions, with an emphasis on frequency of response. This is the same procedure used in the experimental condition during the first learning phase. Following this fluency practice, both groups were given a two week retention period of no practice. A one minute frequency check was done at the end of the two week period. Appendix O contains a schematic of the design components.

Data Analysis

All performance data were recorded on the AC-4 behavior chart (Berquam, 1980). Advantages of this type of chart have been discussed in a previous chapter. Several procedures were used to answer question one (What is the specific nature of the relation between frequency of response and retention?). The first involved analysis of charted data on the AC-4. On individual charts, retention scores were connected to the previous performance score. This provides a measure similar to a

celeration line, although celeration lines generally require repeated measures. Individual lines are then superimposed upon each other on a single chart. This provides a celeration stack, which is a collection of individual trend lines. When a celeration stack is constructed for each of the groups separately, a unique comparison of the retention of the groups can be made. An example of a celeration stack can be found in Appendix D.

From the celeration stacks described above, a procedure similar to analysis of variance can be done. In this procedure, called stratified trend analysis or STRAT analysis (Berquam, in press), individual trend lines are grouped by beginning performance, thus providing a stratified celeration stack. That is, all lines with beginning performance within a specified range, say 50-59 correct per minute, are drawn from a common point at 55 per minute. All lines with beginning performance between 40-49 are drawn from a common point at 45 per minute. This procedure is continued until all lines have been included. Median trend lines from each stratification are then identified and drawn separately. The resulting picture shows the relation between beginning performance and ending performance in a manner that allows proportional differences to be assessed. This procedure is somewhat comparable to analysis of variance, except that median scores are used. An example of STRAT analysis is provided in Appendix E.

To answer question two (Are performance, retention, and celeration scores for the "practiced fluency" condition different from those of the "incidental fluency" condition?), analysis of variance and STRAT analysis

were used. Also, frequency distributions by condition were constructed. Frequency distributions on the the AC-4 are similar to a traditional display, except that the entire display is rotated 90 degrees (see Appendix C). Distributions on the AC-4 have the added advantage that proportional differences can be assessed.

To answer question three (For a given frequency, are there differences in retention between the "practiced fluency" condition and the "incidental fluency" condition?), STRAT analysis and analysis of variance were used.

CHAPTER IV

RESULTS

Descriptive Data on Subjects

Two classrooms of third-graders served as subjects for this study. Both classes were located in the same school building in Orange County, Florida. Generally, the population of this school consisted of students from a low socio-economic area of the County. One classroom served as the experimental group, and the other served as a control, although both groups received the fluency training during the relearning phase of the study. The experimental group had a total enrollment of 21 students. Six students were dropped from the study because of absence or failure to meet a criterion of 90% correct responses on the learning phase of the study. The control group had a total of 22 students, with three dropped due to absence or failure to meet criterion. Thus, the experimental group had 15 students, and the control had 19 students.

Several types of data were analyzed to compare the groups on entering skill levels. In all, six comparisons were made. Raw scores were compared for the communications and mathematics sections of the third grade Florida Statewide Student Assessment Test, a criterion referenced test. Standard scores for total reading, total math, and total battery on the Comprehensive Test of Basic Skills were also

compared. Finally, digit writing rates for numerals 0-9 (in sequence) were compared.

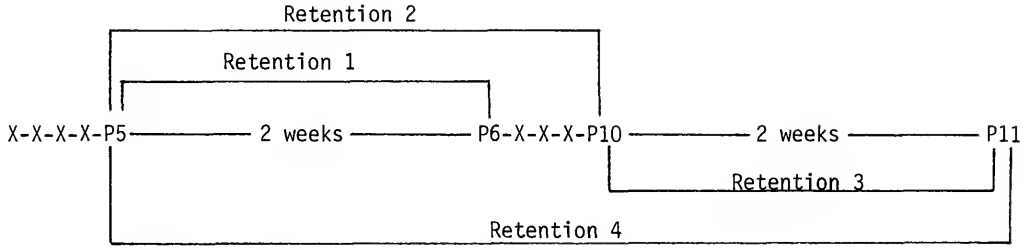
The above data sets were compared using analysis of variance procedures. Five of the comparisons yielded nonsignificant F-ratios; only the math section of the CTBS showed a significant F-ratio, in favor of the experimental group. From these comparisons, it appears that the two classrooms were rather evenly matched. A summary table of the results of these comparisons can be found in Appendix P.

In an attempt to determine if only the better students performed well on the experimental task, Pearson correlation coefficients were calculated between the learning and the retention measures on the experimental task and the school achievement measures listed above that were used for the matching comparisons. A summary of this analysis can be found in Appendix Q.

A moderate correlation was found between performance measures on the experimental task and digit writing rate (.348 to .483). There was also a moderate relation between the experimental task and scores on the communication skills section of the SSAT (.263 to .463). Correlations between the experimental task and the other four measures ranged from -.223 to +.258. Analysis of these data indicates that performance on the experimental task was not highly correlated with school achievement.

General Analysis Procedures

Within the design of the study, four performance measures, all recorded as correct responses per minute, were available for analysis. Figure 1 is a diagram of these scores and the retention measures.



X=Daily Practice
P5=Performance on Day 5 of Learning Phase
P6=Performance after 2 week Retention Phase 1
P10=Performance on Day 5 of Relearning

Retention 1 = P5 to P6
Retention 2 = P5 to P10
Retention 3 = P10 to P11
Retention 4 = P5 to P11

FIGURE 1

DIAGRAM OF RETENTION MEASURES

Performance scores on the fifth and last day of this phase were labeled P5. A two-week retention interval of no practice was then scheduled, followed by five days of daily practice sessions during the relearning phase. Scores on the first day of the relearning phase were labeled P6, and served as a retention check to P5.

Scores on the last day of this phase were labeled P10, and served as an alternate retention check for P5. Following a second two-week retention interval, one measure was taken, and labeled P11. This measure was compared to P10 and also to P5, as retention checks. Four retention comparisons are available from the above four measures. The comparison of scores P5 and P6 was labeled Retention 1; the comparison of scores P5 and P10 was labeled Retention 2; the comparison of scores P10 and P11 was labeled Retention 3, and comparison of scores P5 and P11 was labeled Retention 4.

Scores on P5, the performance measure at the end of the learning phase, were sorted into thirds, thus providing three levels of performance on P5. This was done separately for the two groups, and also for the total combined group. Tables 1 and 2 display scores for the control and experimental groups on P5, P6, P10, and P11, sorted by thirds on P5.

In analyzing the data in this study, no single test or procedure was used as the sole analysis. For example, to answer question one regarding the nature of the relation between frequency of response and retention, several procedures were used. First, correlation coefficients were calculated to determine if there was a relation and to find

TABLE 1

Experimental Group Data, Sorted by P5

	Sub.	P5	P6	P10	P11
Top	105	77	67	71	69
	116	76	59	67	61
	112	76	56	96	80
	103	73	62	73	70
	102	63	50	52	53
Section Totals					
Median		76	59	71	69
Mean		73.0	58.8	71.8	66.6
Middle	104	62	45	74	51
	108	62	17	49	39
	114	60	60	68	61
	111	60	40	68	55
	101	59	31	54	28
Section Totals					
Median		60	40	68	51
Mean		60.6	38.6	62.6	46.8
Bottom	109	58	58	59	52
	107	57	37	72	75
	110	57	23	50	45
	115	48	29	53	X
	113	43	30	44	46
Section Totals					
Median		57	30	53	46
Mean		52.1	35.4	55.6	54.5
Grand Totals					
Median		60	45	67	53
Mean		62.1	44.3	63.3	56.1

TABLE 2

Control Group Data, Sorted by P5

	Sub.	P5	P6	P10	P11
Top	212	66	40	51	64
	217	65	51	70	62
	208	64	40	52	34
	210	63	58	70	55
	215	60	45	52	39
Section Totals					
Median		64	45	52	55
Mean		63.6	46.8	58.0	50.8
Middle	218	59	36	60	57
	209	54	52	56	53
	203	54	46	50	53
	216	54	34	50	34
	204	51	45	50	X
	211	51	19	42	39
	205	50	43	59	49
Section Totals					
Median		54	43	50	49
Mean		53.3	39.3	52.4	47.5
Bottom	214	46	36	54	45
	219	43	32	39	48
	206	42	29	48	40
	207	36	33	34	16
	201	30	27	31	19
	213	29	30	32	33
	202	20	8	17	15
Section Totals					
Median		36	30	34	33
Mean		35.1	27.9	36.4	30.9
Grand Totals					
Median		51	36	50	40
Mean		49.3	38.7	48	41.9

the size and direction of the relation. Second, with the data sorted by levels of P5, several comparisons could be done. Analysis of variance of scores on P6, P10, and P11 was done using levels of P5 for levels of the independent variable. Scores were also compared using the STRAT analysis technique described earlier. This involved analysis of median scores on the AC-4 behavior chart.

With the multiple comparisons being done, a method of summarizing the results was needed. For a particular question or procedure, a matrix was developed. Individual results were classified as showing increasing proportional differences (+), equivalent proportional differences (0), or decreasing proportional differences (-), as described in a previous chapter. Subtotals were then calculated, which provided analysis of results by group, by variable, and by level of P5. Overall totals yield information concerning the generality or reliability of the relation in question.

The STRAT analysis and summary matrix procedure have been used to analyze data related to Question 1, Question 2A, and Question 2B. The summary matrix, while somewhat complex, presents results of all individual contrasts in a way that allows for subtotal analysis for all variables. This procedure is rather unique in this application, but it has added greatly to the analysis of data from the study.

Question 1 Results

The first question posed by this study was "What is the specific nature of the relation between frequency of response and retention?" Three procedures, correlation analysis, analysis of variance, and STRAT analysis were used to analyze data relevant to this question.

Pearson correlation coefficients between the four performance measures were calculated. Results are displayed in Table 3. These figures range from a low of .596 to a high of .832. All were significant at the .01 level.

TABLE 3
Correlations Between Various Performance Measures (Combined Groups)

	P6	P10	P11
P5	.705	.832	.669
P6		.734	.596
P10			.719

(N = 34)

Variables

P5 = Performance at end of learning phase 1

P6 = Performance at end of retention phase 1

P10 = Performance at end of relearning phase

P11 = Performance at end of retention phase 2

The ranking procedure described earlier, using levels of measure P5, was used to provide three levels of classification for the independent variable in the ANOVA and STRAT analysis procedures.

Tables 4 through 6 contain information concerning the analysis of variance procedure, including results of individual contrasts between the levels of P5. Table 4 presents data for P6, the first check on retention. It can be seen that the overall F-ratio is significant. It should also be noted that there is a linear relationship in all three cases between mean scores on P6 and levels of P5. Two of the three contrasts between the individual levels were significant at the .05 level.

TABLE 4

Comparison of Performance on P6, Sorted by Rank on P5

(Low = Low-53; Middle = 54-61; High = 62-High)

	Mean	F-ratio	p	Contrast	p	
High	49.5	8.015	.0016	High-Middle	.181	(0)
Middle	42.0			High-Low	.001	(+)
Low	30.1			Middle-Low	.015	(+)

Table 5 presents ANOVA data for P10. Results are similar to those for P6. The relationship between P10 and levels of P5 is linear in all three comparisons. Two of the three contrasts were significant at the .05 level.

TABLE 5

Comparison of Performance on P10, Sorted by Rank on P5

(Low = Low-53; Middle = 54-61; High = 62-High)

	Mean	F-ratio	p	Contrast	p	
High	65.4	11.931	.0001	High-Middle	.163	(0)
Middle	58.1			High-Low	.000	(+)
Low	41.9			Middle-Low	.001	(+)

Data concerning P11 are presented in Table 6. There is a linear relationship between P11 and P5, and the F-ratio is significant at the .05 level. Two of the three contrasts related to P11 were significant, favoring the higher level of P5.

TABLE 6

Comparison of Performance on P11, Sorted by Rank on P5

(Low = Low-53; Middle = 54-61; High = 62-High)

	Mean	F-ratio	p	Contrast	p	
High	58.0	7.916	.008	High-Middle	.184	(0)
Middle	50.2			High-Low	.001	(+)
Low	35.0			Middle-Low	.017	(+)

The data on Tables 4 through 6 provide an opportunity to examine nine individual contrasts. These contrasts have been labeled with a "+" sign if the results are significant at the .05 level and

favor higher performance on P5. No difference was labeled "0", and if a significant difference favored lower performance on P5, it was labeled "-".

This information was then arranged in a matrix, shown in Table 7. Nine contrasts are included in Table 7. No contrasts favored lower performance on P5. Six contrasts favored higher performance on P5, and 3 showed no difference. Analysis of the subtotals show the contrast between high and low performance and the contrast between middle and low performance on P5 to be quite consistent. The contrast between high and middle performance on P5 shows no difference.

TABLE 7

Summary Matrix of Contrast Results for Analysis of Variance on P6, P10, and P11, by Levels of P5

P6			P10			P11			Totals		
HM	HL	ML	HM	HL	ML	HM	HL	ML	+	0	-
0	+	+	0	+	+	0	+	+	6	3	0

		HM	HL	ML	Grand Totals +	0	-
Totals	+	0	3	3	6	(66.7%)	
x	0	3	0	0		3	(33.3%)
Contrast	-	0	0	0			0 (0%)

For the STRAT analysis, median scores were charted and classified by nature of the relation shown (see Appendix R). Results of these analyses were then summarized in a matrix format on Table 8. Analysis

of subtotals on this matrix indicates that nine contrasts (75%) show increasing proportional differences (+), two contrasts (16.7%) show equivalent proportional differences (0), and one contrast (8.3%) shows decreasing proportional differences (-). Subtotals by contrast show the most consistent results for the High-Middle and High-Low.

TABLE 8

Summary Matrix for STRAT Analysis on Retention Measures

	Ret 1			Ret 2			Ret 3			Ret 4			Totals		
	HM	HL	ML	HM	HL	ML	HM	HL	ML	HM	HL	ML	+	0	-
Medians	+	+	0	+	+	-	0	+	+	+	+	+	9	2	1
Totals	+	0	-	+	0	-	+	0	-	+	0	-			
X Variables	2	1	0	2	0	1	2	1	0	3	0	0			

		HM	HL	ML	Grand Totals +	0	-
Totals	+	3	4	2	9	(75%)	
X	0	1	0	1		2 (16.7%)	
Contrast	-	0	0	1			1 (8.3%)

Contrasts:

HM = High to Middle
HL = High to Low
ML = Middle to Low

Ret 1 = P5 to P6
Ret 2 = P5 to P10
Ret 3 = P10 to P11
Ret 4 = P5 to P11

Summary of Question 1 Results

Data related to Question 1, concerning the nature of the relation between frequency of response and retention, were analyzed in steps. First, Pearson Correlation coefficients were calculated between a performance measure and subsequent retention measure. These figures, ranging from +.596 to +.832, indicate a moderate to strong, positive relation between performance and retention.

Analysis of variance was then done using three levels of P5, the first performance measure, as levels of the independent variable. Results for measures P6, P10, and P11 each yielded significant F-ratios. Contrasts between levels of P5 showed six of nine comparisons to be significant. The contrasts between high and middle levels of P5 were not significant, although results were linear.

To determine the specific nature of the relation, STRAT analysis using median scores was done. Nine of the 12 comparisons (75%) show increasing proportional differences. That is, the greater the performance on P5, the greater is the proportion of performance retained. Only one of twelve contrasts show decreasing proportional differences. Celeration ratios were calculated for each of the contrasts. This quantification of differences between slopes agrees with the STRAT analysis. The high-middle ratio was x1.14, the high-low ratio was x1.21, and the middle-low ratio was x1.06.

Question 2

The first part of question two concerns the effects of specific fluency-building practice on performance and retention. The first

analysis used to answer this question involved the use of frequency distributions of scores of P5, P6, P10, and P11. These distributions are displayed in Appendix S, using the AC-4 behavior chart. Table 9 shows the medians and ranges for the two groups on measure P5, P6, P10, and P11.

TABLE 9

Medians and Ranges for Distribution on P5, P6, P10, and P11, Sorted by Group

	P5	P6	P10	P11
Experimental				
Median	65	45	65	55
High-Low	77-43	67-17	96-44	80-28
Range	x1.79	x3.94	x2.18	x2.86
Control				
Median	55	35	55	45
High-Low	66-20	58-8	70-17	64-15
Range	x3.3	x7.25	x4.12	x4.27

A second technique employed to examine effects of the experimental condition was analysis of variance. The two groups served as levels of the independent variable, and measures P5, P6, P10, and P11, served as dependent variables. Results are listed in Table 10. Analysis of Table 10 shows that on all four performance measures, mean scores for the experimental group are higher than means for the control group. Only the comparison for P6 failed to be significant at the .05 level.

TABLE 10

Comparisons of Group Means on Performance Measures

	P5	P6	P10	P11
Experimental	62.1	44.3	63.3	56.1
Control	49.3	38.7	48.0	41.9
F-ratio	9.663	1.53	10.961	7.281
p-value	0.0039	0.2254	0.0023	0.0113

In addition to the analysis of variance comparisons, median scores were compared using STRAT analysis. Differences at each measure were labeled as showing increasing (+), equivalent (0), or decreasing (-) proportional differences, from the reference point of the experimental group. Charts of individual comparisons in this analysis are contained in Appendix T. A summary of these comparisons is shown in matrix form on Table 11.

TABLE 11

Summary Matrix of STRAT Analysis Comparing Experimental and Control Scores, Sorted by P5

	Ret 1	Ret 2	Ret 3	Ret 4	Totals x Range +	0	-
Top Third	+	+	-	+	3	0	1
Mid Third	-	+	-	-	1	0	3
Bottom Third	-	+	-	-	1	0	3
Entire Range	0	+	0	+	2	2	0

					Grand Totals
Totals	+	1	4	0	2	7 (43.8%)
X	0	1	0	1	0	2 (12.5%)
Variable	-	2	0	3	2	7 (43.8%)

Inspection of Table 11 shows several trends. Comparisons for the top third of each group favor the experimental group, and scores for the middle third and bottom third favor the control group. Comparisons over the entire range of scores favor the experimental group.

Retention 2 (P5 to P10) shows the most consistent trend, in favor of the experimental group. Retention 3 (P10 to P11) favors the control group. Overall, seven of 16 comparisons favor the experimental group, seven favor the control group, and two show no difference.

Question 2B

The second part of Question 2 concerns the differences in retention between experimental and control groups when beginning performance is equated. In an attempt to answer this question, performance on measure P5 was matched for the two groups. This was done in two ways: a) by comparing the top third of the control group (based on P5) to the middle third of the experimental group; and b) by comparing the middle third of the control group to the bottom third of the experimental group. Median scores on measure P5 for these comparisons are listed on Table 12.

TABLE 12

Matched Medians on Measure P5

	Experimental - Control		Experimental - Control	
	(Middle)	(Top)	(Bottom)	(Middle)
Median	60	64	54	57

Comparisons were then made at each retention measure, noting whether the relation showed increasing (+), equivalent (0), or decreasing (-) proportional differences, with respect to the experimental scores. Individual analysis charts are in Appendix U.

A summary of the eight comparisons made concerning this question is contained in a matrix in Table 13. Analysis of this matrix indicates an advantage favoring the control group, when matched on measure P5. Two of eight comparisons favored the experimental group, two showed no difference, and four comparisons favored the control group.

TABLE 13

Summary Matrix of Comparisons Between Control and Experimental Groups, Matched on P5

	Ret 1	Ret 2	Ret 3	Ret 4	Totals x Match +	0	-
Experimental (Middle) Control (Top)	0	+	-	0	1	2	1
Experimental (Bottom) Control (Middle)	(-)	+	(-)	(-)	1	0	3

						Grand Totals
Subtotals	+	0	2	0	0	2 (25%)
X	0	1	0	0	1	2 (25%)
Variable	-	1	0	1	1	4 (50%)

The sign (-) indicates that performance on that measure for the experimental group was actually below that of the control group.

Summary of Question 2 Results

Analysis of variance was used to compare experimental scores to control group scores. This analysis indicates that on all four performance measures (P5, P6, P10, and P11), mean scores for the experimental group were higher than mean scores for the control group. Only the comparison for measure P6 failed to be significant.

STRAT analysis was used to compare median scores on the AC-4 behavior chart. Comparisons were evenly divided, with seven of 16 contrasts favoring increasing proportional differences, and seven indicating decreasing proportional differences, between the experimental and control groups.

For the second part of Question 2, the middle third of the experimental group was compared to the top third of the control group. Also, the bottom third of the experimental group was compared to the middle third of the control group. This procedure nearly matched performances on P5, although in each case, the median for the control group was greater than that for the experimental group. Two of eight comparisons using STRAT analysis favored the experimental group, and four of eight favored the control group.

To meaningfully compare retention measures between the experimental and control groups, the design of the study provided several features that controlled for specific characteristics of "practice effect." The design provided the same number of practice sessions for each group. However, the experimental or "practiced fluency" group worked for only one minute each day. Correct responses per minute were recorded for this group on each of the five days of this phase. The control or

"incidental fluency" group practiced daily, but under no time restraints. They were instructed to complete the entire practice sheet of 100 responses, on each of the first four days. On the fifth day of this phase, the control group did a one-minute performance sample exactly as the experimental group did.

With this design feature, the control group had more total responses, and more time on task. Appendix V contains a chart displaying the frequency distribution for each of the groups on the variables of total responses and time on task. Median scores for these variables are shown in Table 14.

TABLE 14

Median Scores for Time on Task and Total Responses

	Experimental	Control	Ratio
Time on Task			
Median	5 minutes	8.4 minutes	x1.68
Total Number of Responses			
Median	279	450	x1.61

Total time on task for subjects in the control group was calculated by dividing total number of responses by response rate on measure P5. It can be seen from inspection of Table 14 that the control group had 68% more time on task, and 61% more total responses.

Comparison of Retention Before and After Fluency Training

Another measure of the effect of the fluency training can be obtained by comparing the control group retention slopes before the fluency training (P5 to P6) to retention after fluency training (P10 to P11). Analysis of the charts in Appendix R indicates that the slopes were flatter after the fluency training, even though frequency of P10 was not as high as P5.

Celeration ratios were calculated between control and experimental group slopes. The ratio for the Retention 2 measure was greatest (x1.13). The ratio for Retention 3, when the control group also had the benefit of the fluency training, was ÷1.01, indicating essentially parallel lines, or equivalent proportional differences in retention for this measure.

CHAPTER 5

DISCUSSION

Discussion of the results of this study will follow the sequence set forth in the Results Chapter. In the first section, discussion will center around Question 1, "What is the nature of the relation between frequency of response and retention?" The next section will concern the differences in results between the experimental or "practiced fluency" group, and the control or "incidental fluency" group. The third section will discuss the results of the experimental-control comparisons in relation to the "instructional costs" (total time and total responses required) for each group. When it is relevant, results will be related to studies reviewed in Chapter 3. The last section of this chapter will include discussion of possible applications of these findings in an educational setting, and suggestions for further research.

Question 1

The investigation concerning the basic relation between frequency and retention proceeded through several steps. First, Pearson correlation coefficients were calculated between a given measure on the experimental task and all measures that followed it. Coefficients ranged from moderate to strong, and all were positive. The strongest

relation was betwee P5 (the last learning measure) and P10 (the last relearning measure). This correlation was +.832. The strength and consistency of these correlations indicate a strong relation between frequency-based performance and retention measures.

The second stage of this analysis involved sorting scores into thirds, based on measure P5, the last learning measure. This provided three levels of the independent variable that could be used in analyzing the results. Analysis of variance was then done on measures P6, P10, and P11, and all yielded significant overall F-ratios. Planned contrasts between levels of P5 indicated a linear relationship between these levels and each of the other measures, although the p-value for the high-middle contrast was never smaller than .184. This analysis indicates that higher retention is always associated with higher levels of P5, although the differences between the high and middle scores were not statistically significant.

The STRAT analysis procedure provided several comparisons concerning Question 1. First, it supported the results of the analysis of variance procedure, in finding that in all cases, greater retention was associated with higher levels of P5, the last learning measure. This is the same result reported by Krueger (1929, 1930), and led him to conclude that "the higher degree of overlearning always has the higher average retention scores for each interval" (Krueger, 1930, p. 163). Underwood and Keppel, (1963) in summarizing an investigation of retention which varied degree of learning and association level of material, wrote "speaking now in general of the reliable facts of for-

getting, it appears that degree of learning is the only variable involved in a substantial way in retention" (Underwood & Keppel, 1963, p. 16).

Thus, the results pertaining to the first part of Question 1 support the findings of earlier studies. However, the present study attempted to go a step further, to specify more precisely the relation between learning and retention. Using STRAT analysis and the display characteristics of equal-ratio graphs such as the AC-4 behavior chart, proportional differences in retention were analyzed. The basic question is whether retention is a constant percentage of previous performance (shown as parallel lines on the AC-4) or if the percentage retained varies with previous performance.

The chart in Appendix R shows the median scores for each of the the three levels of P5. It can be seen that there is a definite trend toward divergence. On the AC-4, this indicates that as performance on P5 (and later P10) increases, the resulting percentage of retention increases. The summary matrix of this analysis shown in Table 13 indicates that of 12 comparisons (four measures times three contrasts each), nine show proportional retention increasing as previous performance increases. Only one contrast shows a decreasing trend. Analysis of individual contrasts by level of P5 indicates that the high-low contrast was most consistent, with all four measures showing increasing proportional retention. The high-middle contrast shows three of four cases where the proportion of retention increased with P5. The fourth case in this contrast shows equivalent proportional retention.

This result is somewhat different from that shown by the analysis of variance procedure, where the high-middle comparison was not significant. There are several possible explanations. First, while the analysis of variance contrasts were not significant, the means were linear. A second possible explanation could be the influence of "outliers" or extreme scores on the means used in analysis of variance. Third, the STRAT analysis was used to investigate proportional differences, while the analysis of variance procedure examined only absolute differences. Finally, it is possible that the results of the STRAT analysis are not "statistically significant" either, since no parameters for statistical reliability are currently available for this procedure (Berquam, in press).

To summarize results concerning this part of Question 1, it appears that there is a trend toward higher proportional retention with higher levels of previous learning or performance. Three-fourths (9 of 12) of the contrasts indicate this relation. This result agrees with the reanalysis of previous studies by Ammons et al. (1958), Clement (1978), Hellyer (1962), Krueger (1929), Luh (1922), Peterson and Peterson (1959), Slamecka (1959), and Underwood (1954). The most closely related study is that by Clement (1978). In that study, 25 individual data sets were analyzed, involving various academic skills, for students in grades one through six. All data were recorded as frequency of correct responses. Fifteen of the 25 data sets show equivalent proportional retention in relation to previous learning level, and nine sets show increasing proportional retention. A summary matrix of individual contrasts

(Appendix K) shows that 39 of 75 total contrasts indicated increasing proportional differences, 25 equivalent proportional differences and only 11 (15%) showed decreasing proportional differences. Thus, the Clement (1978) data are similar to those of the present study, although somewhat more evenly divided.

A general conclusion regarding the "proportional" nature of retention at this time would best be conservative. There are indications that previous learning or performance may be a controlling variable, but results have not been consistent enough to rule out other possible sources of control.

Question 2; Matching Comparisons

Question 2 concerns the comparison of experimental and control scores. For this comparison to be meaningful, some measure of comparability of the two groups on entering skills must be made. Since the experimental task used unique content, the matching could not be done on beginning performance on this task.

Scores on six measures of school achievement were compared to determine if the groups were well matched. The total reading, total math, and total battery scores of the CTBS were compared. Scores on the communications and math sections of the Florida Statewide Assessment Test were also compared. Finally, digit writing rates were compared.

Inspection of results of these comparisons (see Appendix P) shows only the scores for the math subtest of the CTBS to be significantly different, with the experimental group mean higher than the control

group mean. Probability levels for the other comparisons range from .2441 to .7089. It would appear that the two groups were fairly well matched on the entering skills that were compared. A comparison of average number of correct responses per day for the accuracy checks made during the initial learning stage of the study shows the experimental group averaged 9.7 correct out of ten, and the control group averaged 9.6 correct responses (see Appendix P).

Appendix Q contains correlations between each of the performance measures of the study, and the matching variables. These correlations range from -.223 to +.483. The highest correlation was between measure P5 and digit writing rate. Measure P11 had a moderate correlation with the communication subtest of the State Assessment Test. These correlations indicate that performance on the experimental task was fairly independent of previous school achievement.

Question 2, Part A

Several procedures were used to compare scores between the experimental and control groups. The distributions for scores on measures P5, P6, P10, and P11 indicate that in all cases the median for the experimental or "practiced fluency" group exceeded the median for the control or "incidental fluency" group. Mean scores for the experimental group on these measures were also greater than the means for the control group. Only the difference in scores on measure P6 (the first retention measure) failed to reach statistical significance (.225).

STRAT analysis was used to examine proportional differences between median scores for the two groups on the four retention measures. Examination of the summary matrix for this analysis (Table 13) shows that seven of 14 contrasts favored increasing proportional differences, and seven contrasts show the opposite relation. Comparison of subtotals on this matrix indicates that retention measure 2 (P5 to P6) consistently shows increasing proportional differences (4 of 4 contrasts). Retention measure 3 (P10 to P11) favors decreasing proportional differences in three of the four contrasts. This measure was taken after both groups received the fluency training during the relearning phase of the study.

Comparison of subtotals by range of performance on P5 indicates most consistent results favoring the experimental group for the top third. The middle and bottom thirds favored the control group. Overall, this analysis indicates that results related to proportional differences between medians of the two groups are rather evenly divided. The differences that favored the experimental group occurred before the control group scores were affected by the fluency training.

Another measure of the relation between fluency training and retention can be obtained by comparing retention from before and after the fluency training. The design of the study provides this comparison for the control group. Analysis of the retention slopes for the control group indicates that the percentage of retention was higher (the slopes were flatter) for the second retention interval (P10 to P11) than for the first interval (P5 to P6). This relation is true for each of the

tri-sections of the control group distribution. Interestingly, for all three sections, measure P10 was less than its related measure, P5, yet the retention after P10 was proportionally greater than that after P5. This same relation was true for the experimental group, although it was not as dramatic. Apparently the fluency training continues to enhance retention, even in relearning conditions.

Question 2, Part B

Another comparison of experimental and control scores was done using STRAT analysis. In this procedure, the middle third of the experimental group (based on levels of P5) was compared to the top third of the control group, and the bottom third of the experimental group was compared to the middle third of the control group. In this way, scores on measure P5 were more evenly matched between the groups. As was shown in Table 12, median scores for the control group actually exceeded those for the experimental group with this matching procedure. The summary matrix for these comparisons (Table 13) shows results rather evenly divided, although retention measure 2 (P5 to P10) again consistently favors the experimental group. Basically, the above analysis indicates that when matched on measure P5, there is little difference between groups on later measures. It must be pointed out, however, that the matching procedures compared different sections of the distributions for the two groups. It is possible that with initial fluency training, the top third of the control group would have matched the top third of the experimental group, rather than the middle third of that group.

"Cost" of Training Procedures

One last point must be made concerning the experimental-control comparisons. The design of the study allowed more total responses and more total practice time for the control group. In fact, median scores for the control group on these measures exceed those of the experimental group by 60-70%. Therefore, the performance and retention measures of the experimental group were achieved at a substantial savings in both time and effort. Even if actual scores for the two groups were comparable, the fluency training proved to be more efficient at producing those scores than was the unpaced or "incidental fluency" method. When we consider that scores for the experimental group exceeded those of the control group, even after the control group received fluency training during the relearning phase, there is additional support for the use of fluency training procedures to enhance performance and retention.

Educational Applications

A very basic finding of this study, which confirmed results of earlier investigations, is that level of previous learning has a very strong relation to subsequent retention of skills. Since this study used frequency of response data, the results indicate that the basic learning/retention relation is true with frequency data as well as with the accuracy based measures used in previous studies.

In an earlier section, a distinction was made between procedures that provide extra practice, and those that specifically increase performance. In this study, the control or "incidental fluency" group

received extra practice, with no emphasis on increasing performance (frequency of correct responses). The experimental group specifically worked to increase performance. In all cases, the experimental group scores exceeded the control group scores. This indicates that the fluency training, which was designed to increase performance, was more effective than merely providing extra practice. When the control group received the fluency training during the relearning phase, subsequent retention for that group was greater than when practice alone was given.

When the actual differences in scores are viewed from the perspective that the higher experimental scores were achieved at a considerable savings in time and effort, the efficiency of the fluency training procedure is emphasized. It seems likely that teachers could increase their students' skill levels, and subsequent retention by using short, concentrated periods of fluency training, similar to those used in this study and those used by precision teachers. This would help promote efficient learning. In describing efficient learning, Robert Gagne (1965) says that "the greatest change in the student's behavior will occur in the shortest period of time"(1965, p. 241). This would seem to be a legitimate goal for education.

Implications for Further Research

Several major implications for further investigation are suggested by results of this study. First, the relation between performance and the proportion of retention has not been definitively established. This phenomenon appears with some regularity, both in the data from this

study and in data from previous work. However, it seems apparent that there are other variables controlling the slope of the retentions, also. These variables need to be identified.

A second area of need involves more applied work. Functional levels of fluency and retention for specific skills need to be identified. This information would aid researchers in the field of precision teaching who are working to define frequency-based performance aims. It would also greatly aid teachers in planning instruction and skill maintenance work with students. This could help prevent and/or remediate loss of performance during periods of no practice, such as over the summer months.

A related question is whether the slope of retention lines is a characteristic of individual students. We must view this work with caution, however. Vergason (1964) helped dispel the notion that retarded individuals always have poorer memory. He found that performance level is a stronger controlling variable than intelligence level. Still, the question has merit. But in the search for variables which control retention, researchers must maintain a scientific perspective, and let empirical data guide their decisions.

One last suggestion concerns the type of task or skill used in the study of retention. With the wide variety of tasks used, there is little evidence yet that retention is task-specific. However, as our questions and answers about retention become more sophisticated, the task variable may be found to influence retention.

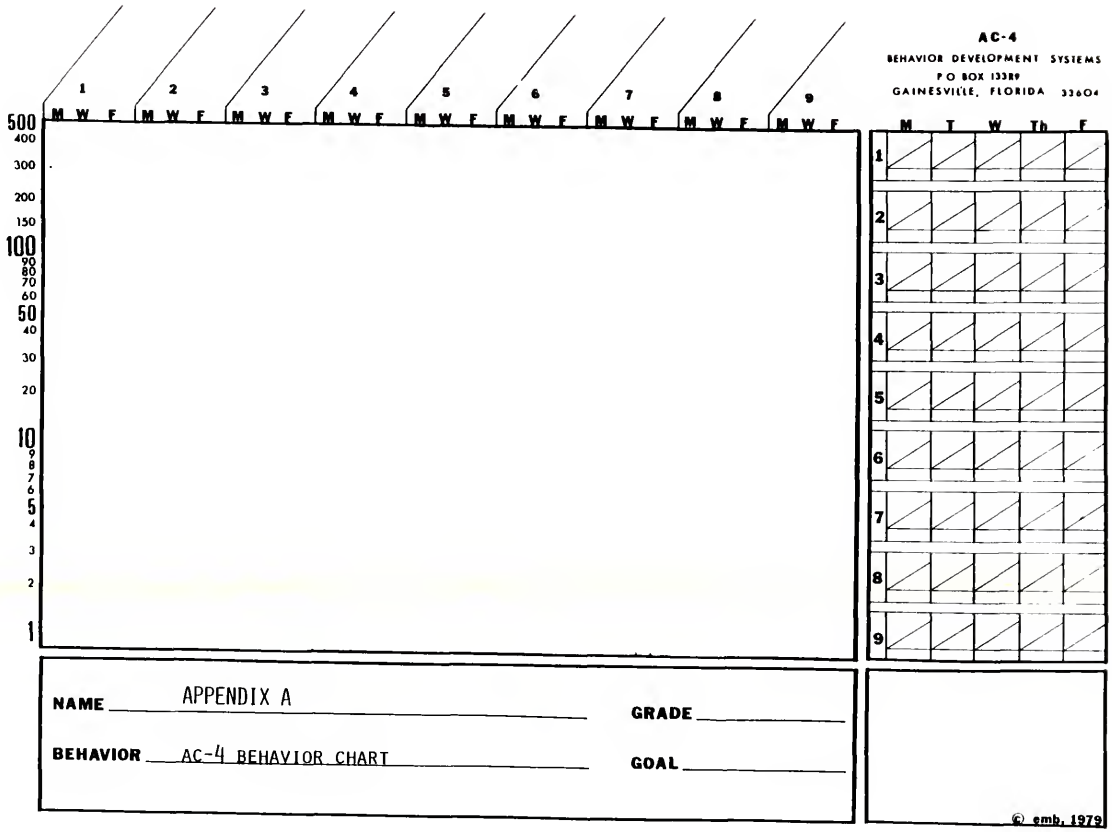
In the late 1800's, Hermann Ebbinghaus initiated what has become a century long search for the nature of retention. In his work, Ebbinghaus insisted upon using scientific methods. He wrote:

> The method of obtaining exact measurements--i.e., numerically exact ones--of the inner structure of causal relations is, by virtue of its nature, of general validity. This method, indeed, has been so exclusively used and so fully worked out by the natural sciences that, as a rule, it is defined as something peculiar to them, as the method of natural science. To repeat, however, its logical nature makes it generally applicable to all spheres of existence and phenomena. Moreover, the possibility of defining accurately and exactly the actual behavior of any process whatever, and thereby of giving a reliable basis for the direct comprehension of its connections depends above all upon the possibility of applying this method. (Ebbinghaus, 1885, p.7)

Certainly this guidance is no less valuable today than it was a century ago.

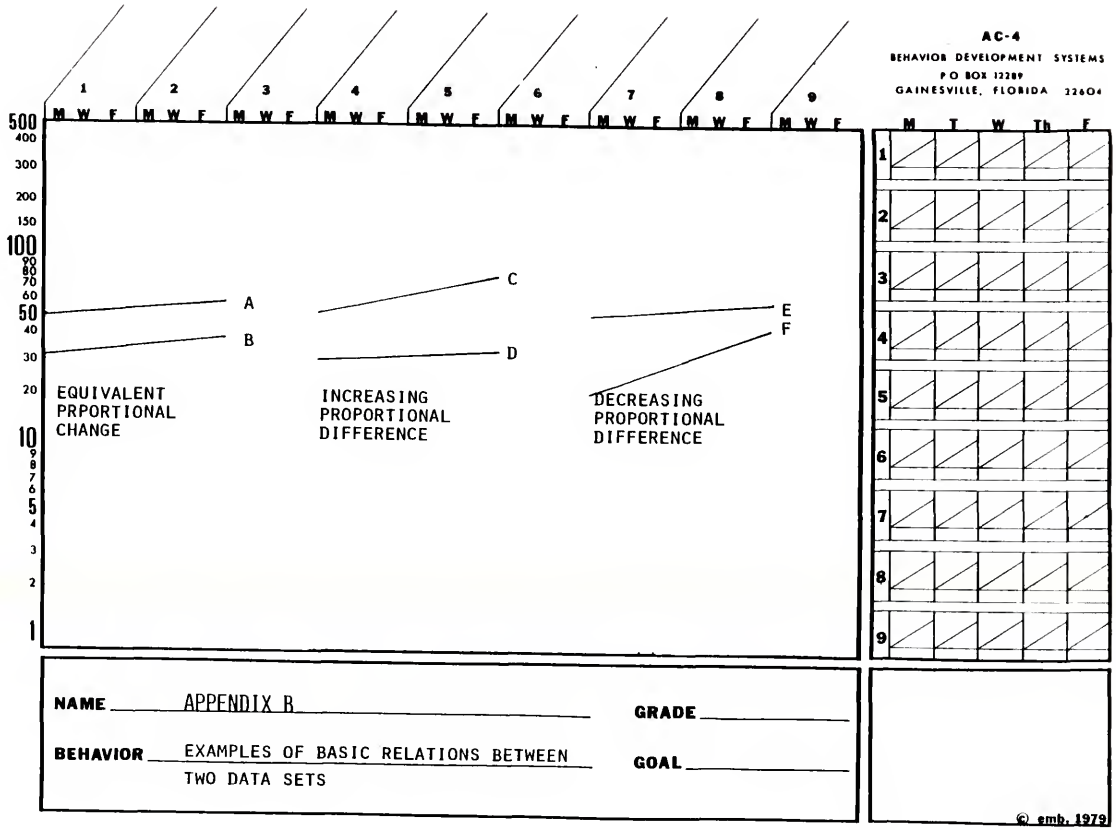
APPENDIX A

AC-4 BEHAVIOR CHART


AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P O BOX 13389
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500 400 300 200 150 100 90 80 70 60 50 40 30 20 10 9 8 7 6 5 4 3 2 1
M T W Th F
1 2 3 4 5 6 7 8 9
NAME APPENDIX A
GRADE
BEHAVIOR AC-4 BEHAVIOR CHART
GOAL
© emb, 1979

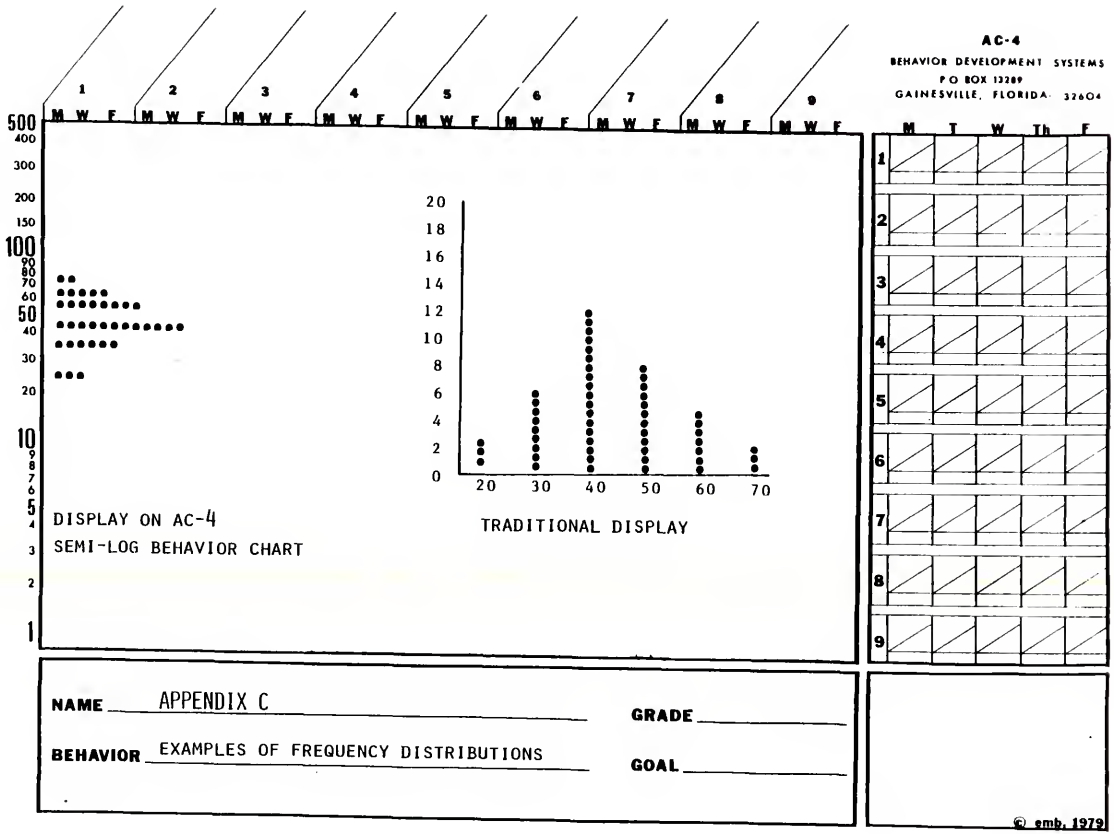
APPENDIX B

EXAMPLES OF BASIC RELATIONS BETWEEN TWO DATA SETS


AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P O BOX 12289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500 400 300 200 150 100 90 80 70 60 50 40 30 20 10 9 8 7 6 5 4 3 2 1
A
B
C
D
E
F
EQUIVALENT PRPORTIONAL CHANGE
INCREASING PROPORTIONAL DIFFERENCE
DECREASING PROPORTIONAL DIFFERENCE
M T W Th F
1 2 3 4 5 6 7 8 9
NAME APPENDIX B
GRADE
BEHAVIOR EXAMPLES OF BASIC RELATIONS BETWEEN TWO DATA SETS
GOAL
© emb. 1979

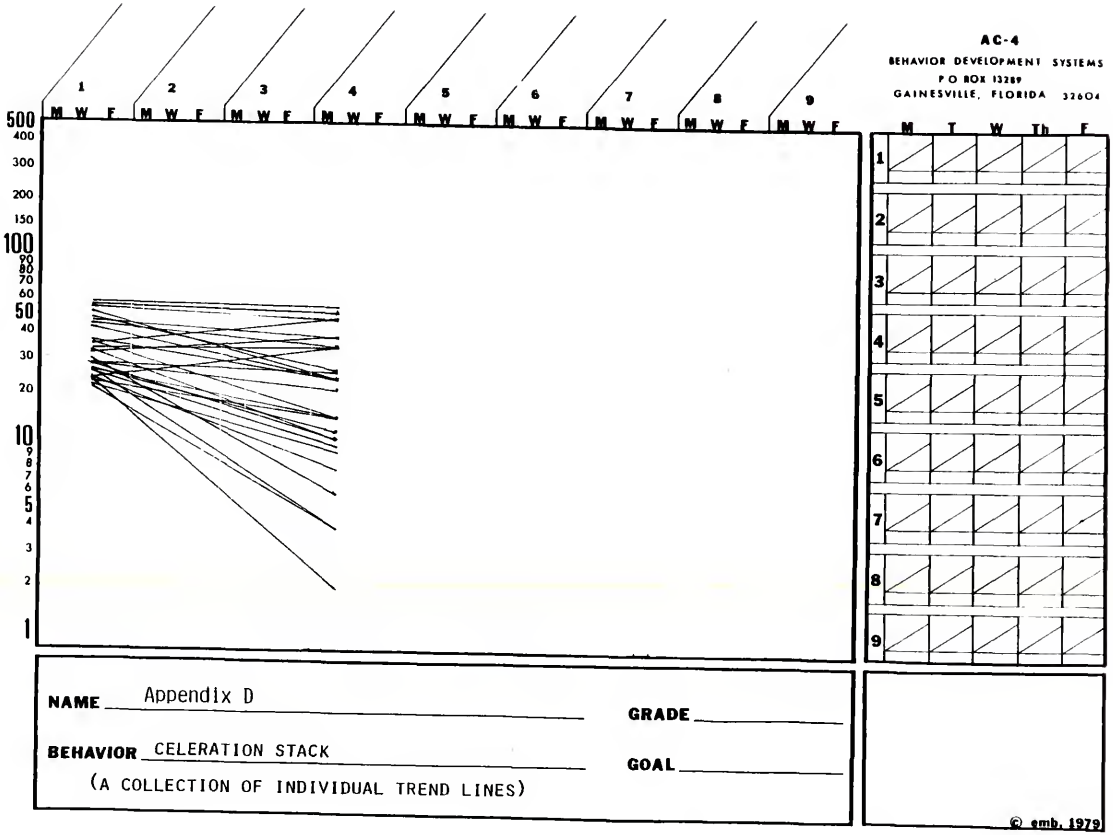
APPENDIX C

EXAMPLE OF FREQUENCY DISTRIBUTION ON AC-4


AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P O BOX 13289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500 400 300 200 150 100 90 80 70 60 50 40 30 20 10 9 8 7 6 5 4 3 2 1
20 18 16 14 12 10 8 6 4 2 0
20 30 40 50 60 70
TRADITIONAL DISPLAY
DISPLAY ON AC-4
SEMI-LOG BEHAVIOR CHART
M T W Th F
1 2 3 4 5 6 7 8 9
NAME APPENDIX C
GRADE
BEHAVIOR EXAMPLES OF FREQUENCY DISTRIBUTIONS
GOAL
© emb, 1979

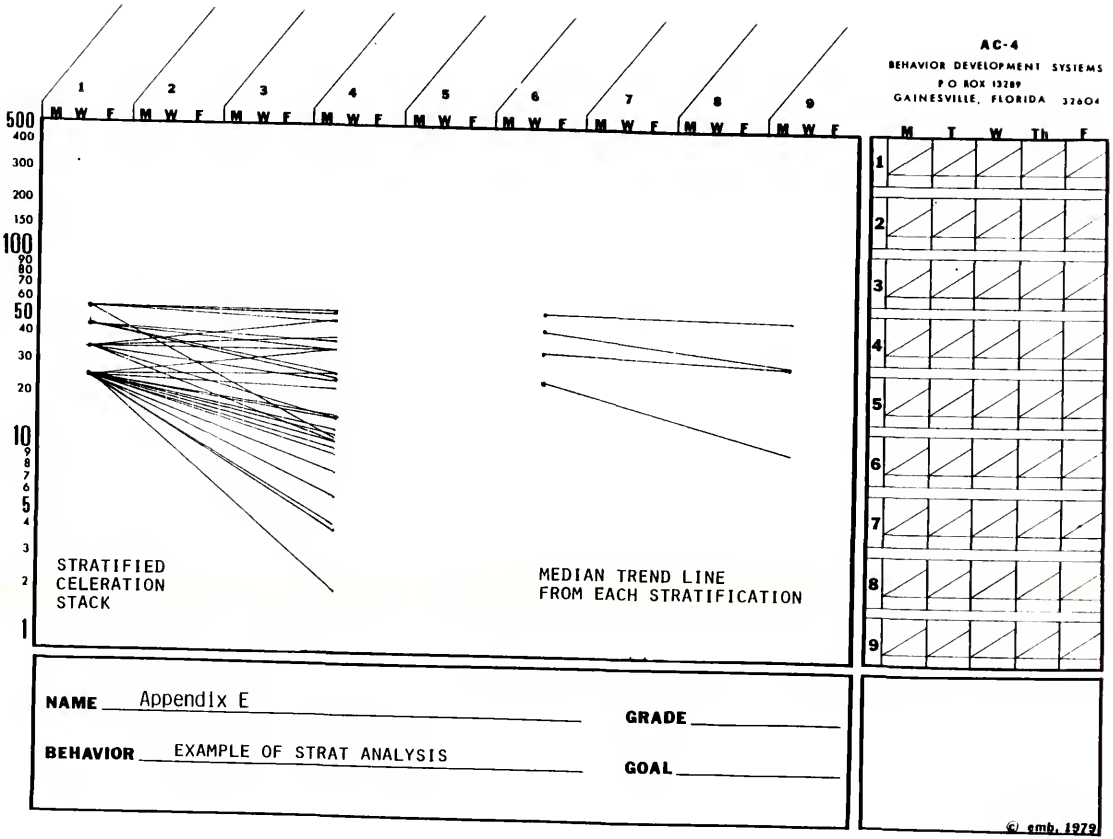
APPENDIX D

EXAMPLE OF CELERATION STACK ON AC-4


AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P O BOX 13289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500 400 300 200 150 100 90 80 70 60 50 40 30 20 10 9 8 7 6 5 4 3 2 1
M T W Th F
1 2 3 4 5 6 7 8 9
NAME Appendix D
GRADE
BEHAVIOR CELERATION STACK
GOAL
(A COLLECTION OF INDIVIDUAL TREND LINES)
© emb. 1979

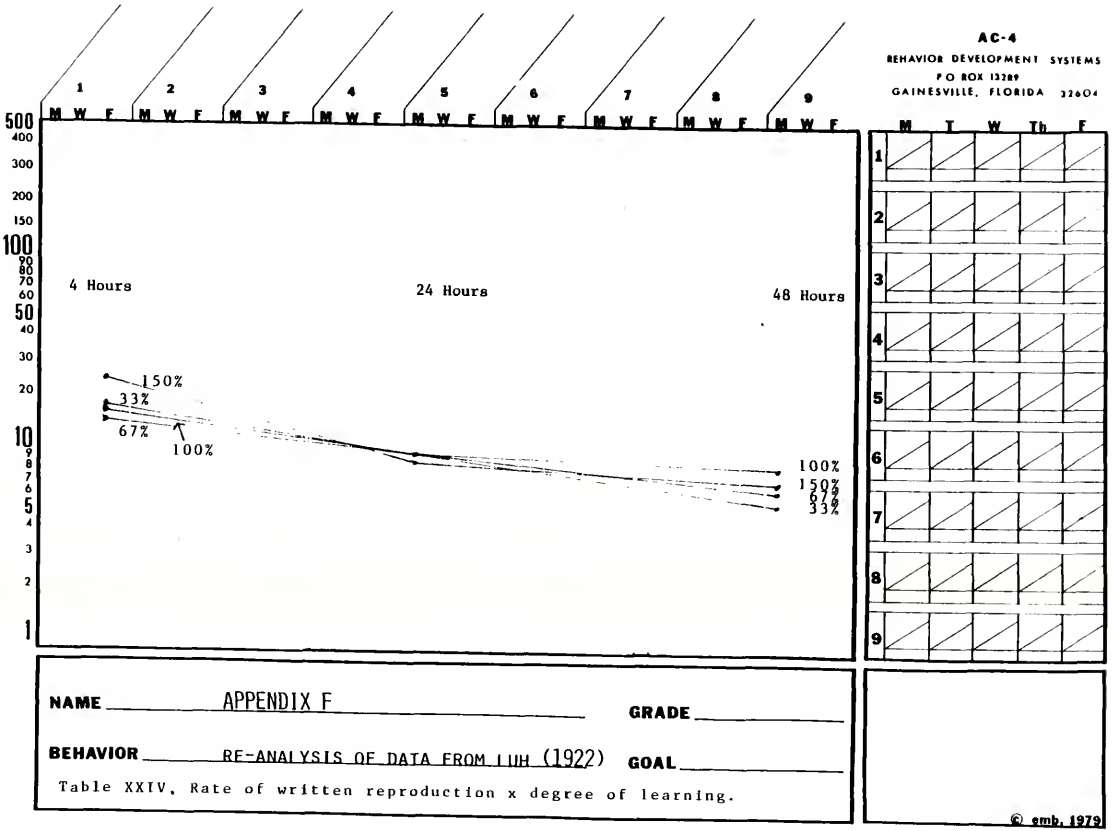
APPENDIX E

EXAMPLE OF STRAT ANALYSIS


AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P O BOX 13289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500 400 300 200 150 100 90 80 70 60 50 40 30 20 10 9 8 7 6 5 4 3 2 1
STRATIFIED CELERATION STACK
MEDIAN TREND LINE FROM EACH STRATIFICATION
M T W Th F
1 2 3 4 5 6 7 8 9
NAME Appendix E
GRADE
BEHAVIOR EXAMPLE OF STRAT ANALYSIS
GOAL
© emb, 1979

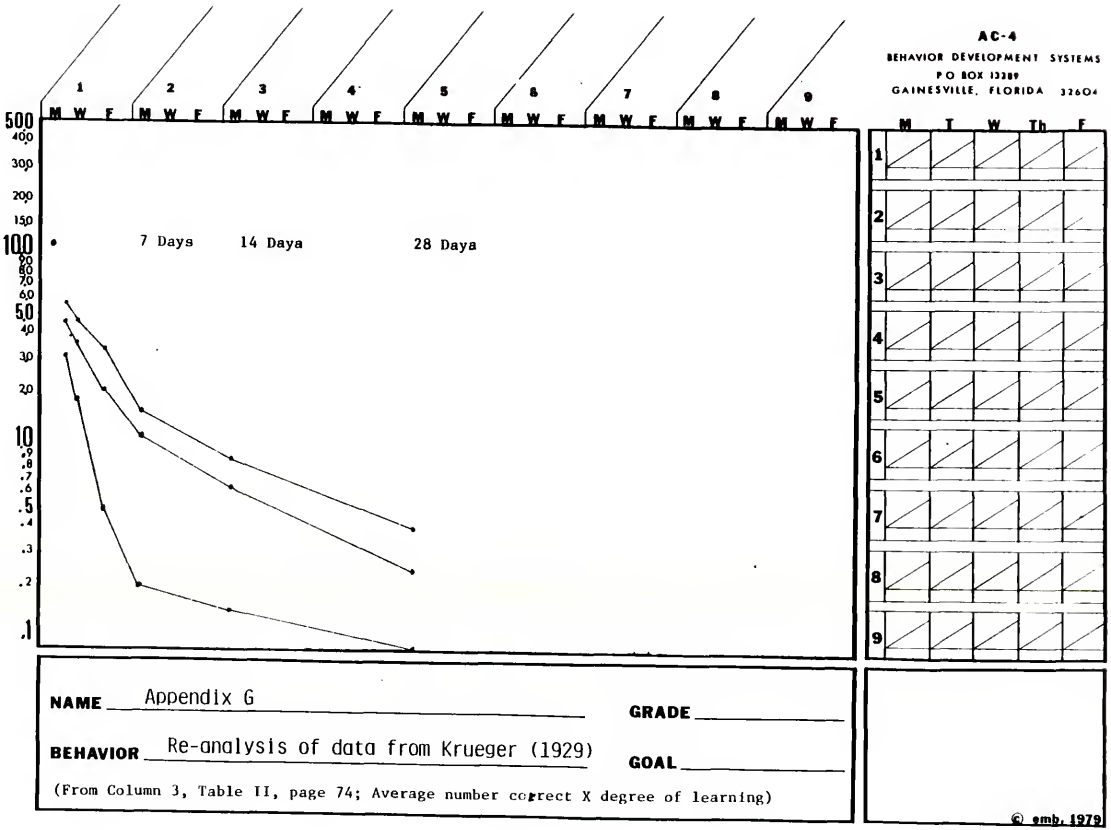
APPENDIX F

RE-ANALYSIS OF DATA FROM LUH (1922)


AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P O BOX 13289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500 400 300 200 150 100 90 80 70 60 50 40 30 20 10 9 8 7 6 5 4 3 2 1
4 Hours
24 Hours
48 Hours
150%
33%
67%
100%
100%
150%
67%
33%
M T W Th F
1 2 3 4 5 6 7 8 9
NAME APPENDIX F
GRADE
BEHAVIOR RE-ANALYSIS OF DATA FROM LUH (1922)
GOAL
Table XXIV. Rate of written reproduction x degree of learning.
© emb. 1979

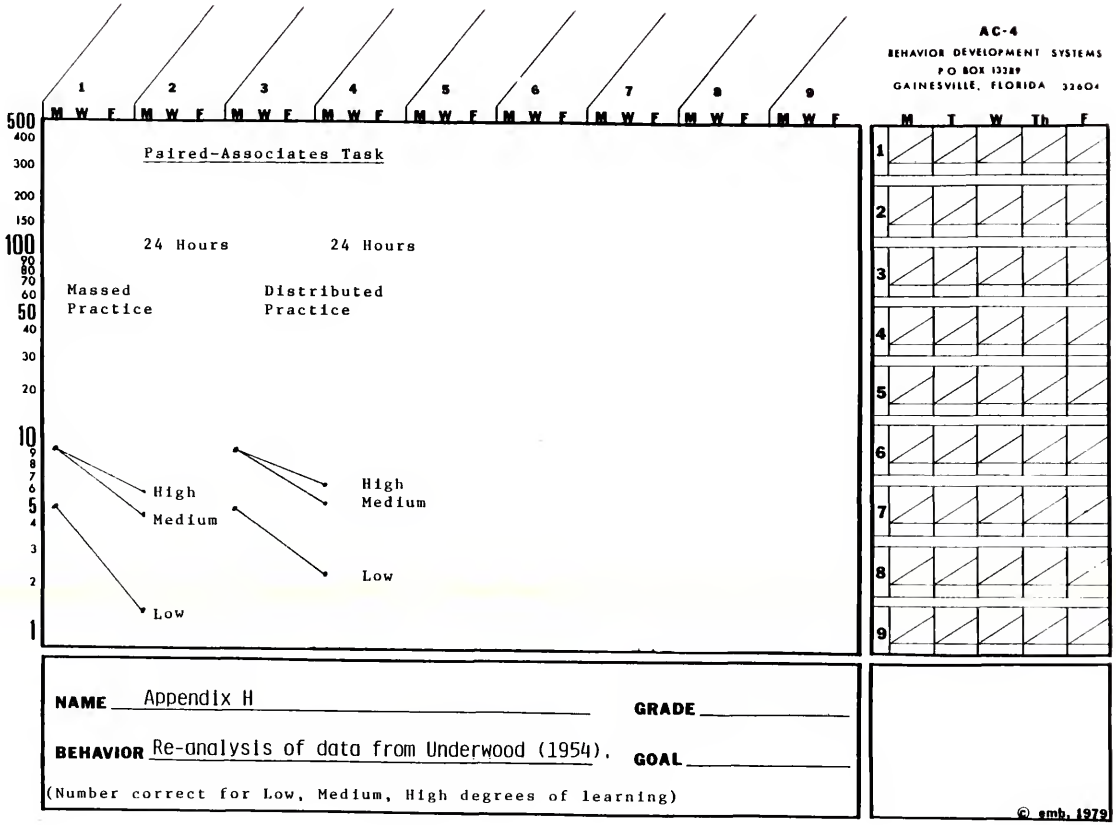
APPENDIX G

RE-ANALYSIS OF DATA FROM KRUEGER (1929)


AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P.O. BOX 13389
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500
400
300
200
150
100
90
80
70
60
50
40
30
20
10
.9
.8
.7
.6
.5
.4
.3
.2
.1
7 Days
14 Days
28 Days
M T W Th F
1
2
3
4
5
6
7
8
9
NAME Appendix G
GRADE
BEHAVIOR Re-analysis of data from Krueger (1929)
GOAL
(From Column 3, Table II, page 74; Average number correct X degree of learning)
© emb, 1979

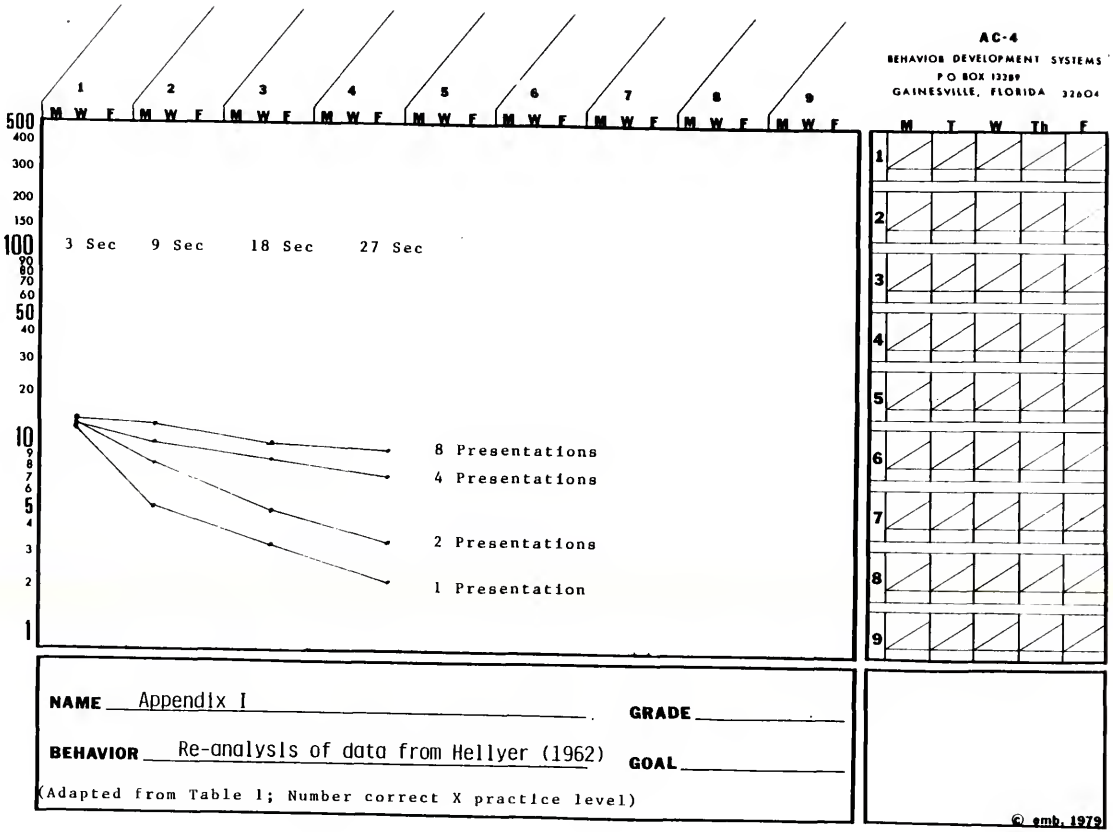
APPENDIX H

RE-ANALYSIS OF DATA FROM UNDERWOOD (1954)


AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P O BOX 13289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500
400
300
200
150
100
90
80
70
60
50
40
30
20
10
9
8
7
6
5
4
3
2
1
Paired-Associates Task
24 Hours
24 Hours
Massed Practice
Distributed Practice
High
Medium
Low
High
Medium
Low
M T W Th F
1
2
3
4
5
6
7
8
9
NAME Appendix H
GRADE
BEHAVIOR Re-analysis of data from Underwood (1954).
GOAL
(Number correct for Low, Medium, High degrees of learning)
© emb, 1979

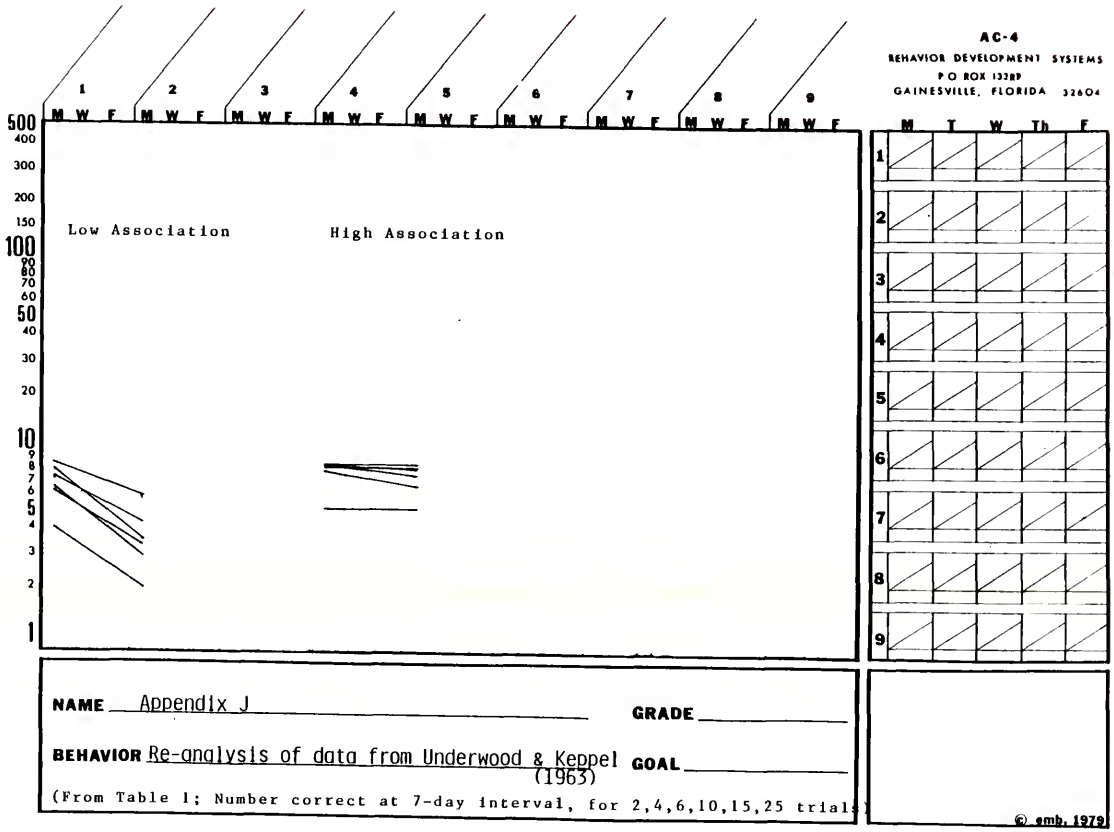
APPENDIX I

RE-ANALYSIS OF DATA FROM HELLYER (1962)


AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P O BOX 13289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500 400 300 200 150 100 90 80 70 60 50 40 30 20 10 9 8 7 6 5 4 3 2 1
3 Sec 9 Sec 18 Sec 27 Sec
8 Presentations
4 Presentations
2 Presentations
1 Presentation
M T W Th F
1 2 3 4 5 6 7 8 9
NAME Appendix I
GRADE
BEHAVIOR Re-analysis of data from Hellyer (1962)
GOAL
(Adapted from Table 1; Number correct X practice level)
© emb, 1979

APPENDIX J

RE-ANALYSIS OF DATA FROM UNDERWOOD & KEPPEL (1963)


1 2 3 4 5 6 7 8 9
M W F
500
400
300
200
150
100
90
80
70
60
50
40
30
20
10
9
8
7
6
5
4
3
2
1
Low Association
High Association
AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P O BOX 13289
GAINESVILLE, FLORIDA 32604
M T W Th F
1
2
3
4
5
6
7
8
9
NAME Appendix J
GRADE
BEHAVIOR Re-analysis of data from Underwood & Keppel (1963)
GOAL
(From Table 1; Number correct at 7-day interval, for 2,4,6,10,15,25 trials)
© amb, 1979

APPENDIX K

Re-analysis of data from Clement (1978)

Appendix K contains a charted re-analysis of retention data from Clement (1978). In this study, frequency of response data were gathered on the following skills: oral reading rate (ORR); silent reading rate (SRR); write numbers 0-9, random (WNR); mixed addition and subtraction facts (A/S); and mixed multiplication and division facts (M/D).

A performance sample was taken in May and again in September, on students in first through sixth grade. Charted scores are the first quartile, mean, and third quartile scores for each grade level.

SUMMARY MATRIX OF SPRING-FALL RETENTION DATA (CLEMENT, 1978)

	ORR			SRR			WNR			A/S			M/D			Totals x Grade		
	HM	HL	ML	HM	HL	ML	HM	HL	ML	HM	HL	ML	HM	HL	ML	+	0	-
1st	0	+	+				+	+	+	0	-	-				5	2	2
2nd	0	-	0	+	+	+	0	0	0	+	+	+				6	5	1
3rd	+	+	+	0	0	0	0	0	0	0	+	+	-	-	-	5	7	3
4th	+	+	+	0	0	0	+	+	+	+	+	+	+	+	+	12	3	0
5th	0	0	0	0	0	0	-	-	0	+	+	0	+	+	+	5	8	2
6th							-	-	0	+	+	+	+	+	+	6	1	2

Totals	+	8	3	3	14	9
x	0	6	9	11	2	0
Variable	-	1	0	4	2	3

		HM	HL	ML	GRAND TOTALS +	0	-
Totals	+	12	14	13	39 (52%)		
x	0	10	6	10		26 (34.6%)	
Contrast	-	3	5	2			10(13.3%)

ORR=Oral Reading Rate
SRR=Silent Reading Rate
WNR=Write Numbers Random
A/S=Mixed Add-Sub. Facts
M/D=Mixed Mult.-Divide Facts

(+)=Increasing Proportional Differences
(0)=Equivalent Proportional Differences
(-)=Decreasing Proportional Differences

APPENDIX L

LITERATURE REVIEW SUMMARY TABLES

SUMMARY OF VERBAL LEARNING STUDIES

STUDY	SUBJECTS	TASK	INTERVAL	DEGREES	RESULTS
Ebbinghaus, 1885	Himself	Lists of 16 non-sense syllables	24 hours	8,16,24,32, 42,53,64 repetions	Retention increased in proportion with repetions.
Luh, 1922	10 college students & faculty	Lists of 12 non-sense syllables	20 minutes-48 hours	33%,67%, 100%,150% learning	Performance increased with practice; retention proportionally greater with higher performance.
Krueger, 1929	120 college students	Lists of 12 non-sense syllables	1,2,4,7,14 28 days	100%.150%, 200% learning	Retention proportionally greater for higher learning levels.
McGeoch, 1929	10 college faculty & students	Lists of non-sense syllables	5 minutes	6,11,16,21 repetitions	Repetition improved recall and relearning.
Underwood, 1954	120 college students	Serial lists, P-A lists	24 hours	50%,100%, 800% learning	Performance increased with practice, up to 100% level; retention proportionally greater for higher learning levels.
Richardson, 1956	60 college students	concept words	24 hours	8, 12, 16 trials	Inhibition decreased, retention unchanged with extra trials.

SUMMARY OF VERBAL LEARNING STUDIES

STUDY	SUBJECTS	TASK	INTERVAL	DEGREES	RESULTS
Briggs, 1957	80 college students	10 paired-associates	20 minutes	2, 5, 10, 20 trials	Performance and retention increased, inhibition decreased with extra trials.
Gilbert, 1957	27 soldiers	Recall facts from narrated story	15 minutes, 24,48 hours	100%,200%, 300% learning	Retention increased proportionally with learning level.
Rock, 1958	Not given	List of nonsense syllables	Not given	Not given	Repetition improved retention.
Estes et al., 1960	48 college students	8 paired-associates	Immediate	Not given	Retention increased with extra trials.
Slamecka, 1959	140 college students	Recall facts from prose	5 minutes	1,3,5,7,9 repetitions	Repetition increased performance. Retention proportionally greater for higher performance level.
Peterson & Peterson, 1959	48 college students	Trigrams	3,9,18 seconds	0,1,3 seconds rehearsal	Retention proportionally greater with more rehearsal.
Hellyer, 1962	25 housewives	Trigrams	3,9,18,27 seconds	1,2,4,8, repetitions	Retention proportionally greater with more repetitions.

SUMMARY OF VERBAL LEARNING STUDIES

STUDY	SUBJECTS	TASK	INTERVAL	DEGREES	RESULTS
Wollen, 1962	136 college students	Paired-associates	Immediate, delayed	Not given	Repetition increased retention.
Postman, 1962	Not given	2-syllable items	7 days	100%, 150%, 200% learning	Retention increased with learning.
Underwood & Keppel, 1963	432 college students	Paired-associates	1 day, 7 days	2,4,6,10,15 25 trials	Performance increased with trials, up to 10 trials; some proportionally greater, some equivalent retention curves.
Reynolds & Glaser, 1964	75 high school students	11 science terms	2 days, 3 weeks	50%, 100%, 150% learning	No differences in retention due to practice.
Vergason, 1964	64 normal, 64 retarded students	Paired-associates	1 day, 30 days	1 correct, 5 correct trials	Retention increased with practice; no differences in learning scores or 30-day retention due to IQ.
Lance, 1965	64 normal, 64 retarded students	Low & high association paired-associates	30 days	1 correct, 5 correct trials	Retention increased with practice; no differences in savings scores due to IQ.
Kratochwill et al., 1971	64 pre-school students	4 common nouns	1 hour	2 correct trials; 10 & 20 extra trials	Linear relation between practice and retention; significant difference only for 20 trial level.

SUMMARY OF VERBAL LEARNING STUDIES

STUDY	SUBJECTS	TASK	INTERVAL	DEGREES	RESULTS
Clement, 1978	Over 400 elementary students	Reading & math tasks	4 months	Based on quartiles of distributions of rate data	52% of contrasts showed increasing proportional retention due to learning level; 34.6% showed equivalent, 13.3% showed decreasing proportional retention due to learning level.
		* * * * * * * * * * * * * * * *			
Roberts, 1972	10 pigeons	Delayed matching to sample	0,1,3,6, seconds	FR1,FR5, FR15 (1,5, 15 repetitions)	Retention increased in proportion to repetitions.
Brehmer & Lindberg, 1973	140 college students	Single-cue probability learning	1 day, 7 days	100%,200%, 300% learning	No differences due to degree of learning.

SUMMARY OF MOTOR LEARNING STUDIES

STUDY	SUBJECTS	TASK	INTERVAL	DEGREES	RESULTS
Krueger, 1930	192 college students	Trace a finger maze	1,2,4,7,14, 28 days	100%, 150%, 200%,300% learning	Retention increased in proportion with learning level
Rubin-Rabson, 1941a	9 piano teachers	Short music passages	2 weeks, 7 months	100%,200%, learning	Small differences after 2 weeks; no differences after 7 months.
Rubin-Rabson, 1941b	9 piano teachers	Short music passages	2 weeks, 7 months	100%,150%, 200%.300% learning	No differences.
Buxton, 1943	276 college students	Pursuit-rotor	10 minutes	90,360,630 seconds on target	Performance increased, reminiscence decreased with practice.
Irion, 1949	75 college students	Pursuit-rotor	5 minutes	10,20,30 40, trials	Performance increased, reminiscence decreased with practice.
Jahnke, 1958	240 high school students	Pursuit-rotor	10 minutes, 1 day,1 week	1, 2½, 5, 10 minutes of practice	Performance increased with practice; retention increased in proportion with practice.
Ammons et al., 1958	538 college students	Complex sequential-motor task	1 minute, 1 day; 6, 12,24 months	5 trials, 30 trials	Performance increased with practice; proportional re- greater for high performance group.

SUMMARY OF MOTOR LEARNING STUDIES

STUDY	SUBJECTS	TASK	INTERVAL	DEGREES	RESULTS
Fleishman & Parker, 1962	62 college faculty	Complex visual-motor task	1, 5, 9, 14, 24 months	Based on performance	Retention increased with performance level; similar relearning scores.
Purdy & Lockhart, 1962	36 college women	Toss-catch, balance tasks	1 year	High, Average, Low	Retention increased in proportion with performance.
Hammerton, 1963	18 college faculty	Complex visual-motor task	6 months	2 levels, based on performance	Retention greater for high performance group; relearning scores similar.
Chasey, 1971	98 retarded students	Diagonal hop	4 weeks	1,3 successful trials	Relearning scores favored overlearning.
Melnick, 1971	80 college students	Stabilometer	1 week, 1 month	100%, 150%, 200%, 300% learning	Retention significantly higher for 300% level.
Chasey & Knowles, 1973	90 mild, moderate, severe retarded boys	Bean bag toss	5 weeks	3 or 9 criterion trials	Retention increased with practice; advantage of effect increased with level of retardation.
Chasey, 1977	100 retarded boys	Stabilometer	8 weeks	100%, 150%, 200%, 250% learning	Retention increased with learning.

LITERATURE REVIEW SUMMARY TABLE

Total number of studies reviewed	37
Total number of studies reviewed with results favoring overlearning or extra practice.	29
Total number of studies reviewed with results indicating no differences	4
Total number of studies reviewed showing mixed results	4
Total number of studies reviewed which included performance change measures,	11
Total number of performance change studies with results favoring higher performance	11
Total number of studies reviewed with results indicating increasing proportional retention due to practice or performance level.	7
Total number of studies reviewed with results indicating equivalent proportional retention due to practice or performance level.	7
Total number of studies reviewed with results indicating decreasing proportional retention due to practice or performance level.	0
Total number of studies reviewed with results indicating both increasing and equivalent proportional retention.	2

APPENDIX M

EXPERIMENTAL TASK PRACTICE SHEETS

GAC 1	MUN 2	NUR 3	SOQ 4	TEZ 5	TOF 6	KAR 7	DUP 8	DAK 9	LAH 0
MUN	SOQ	GAC	NUR	TEZ	KAR	DUP	TOF	LAH	DAK
TOF	KAR	NUR	SOQ	MUN	GAC	TEZ	DUP	DAK	LAH
TEZ	NUR	DAK	LAH	GAC	TOF	DUP	SOQ	KAR	MUN
GAC	NUR	TOF	LAH	MUN	SOQ	DUP	KAR	LAH	DAK
KAR	MUN	DUP	SOQ	TOF	GAC	LAH	NUR	DAK	TEZ
SOQ	NUR	KAR	TOF	DAK	MUN	GAC	TEZ	DUP	LAH
DAK	LAH	TOF	DUP	KAR	TEZ	NUR	GAC	SOQ	MUN
TOF	GAC	DUP	SOQ	LAH	NUR	DAK	TEZ	MUN	DAK
SOQ	LAH	TEZ	DUP	GAC	MUN	TOF	KAR	NUR	DAK
MUN	DAK	LAH	SOQ	TOF	GAC	DUP	NUR	KAR	TEZ

SOQ	NUR	KAR	TOF	DAK	MUN	GAC	TEZ	DUP	LAH
DAK	LAH	TOF	DUP	KAR	TEZ	NUR	GAC	SOQ	MUN
TOF	GAC	DUP	SOQ	LAH	NUR	DAK	TEZ	MUN	DAK
SOQ	LAH	TEZ	DUP	GAC	MUN	TOF	KAR	NUR	DAK
MUN	DAK	LAH	SOQ	TOF	GAC	DUP	NUR	KAR	TEZ
MUN	SOQ	GAC	NUR	TEZ	KAR	DUP	TOF	LAH	DAK
TOF	KAR	NUR	SOQ	MUN	GAC	TEZ	DUP	DAK	LAH
TEZ	NUR	DAK	LAH	GAC	TOF	DUP	SOQ	KAR	MUN
GAC	NUR	TOF	LAH	MUN	SOQ	DUP	KAR	LAH	DAK
KAR	MUN	DUP	SOQ	TOF	GAC	LAH	NUR	DAK	TEZ

APPENDIX N

CORRELATION COEFFICIENTS BETWEEN A SHORT TERM MEMORY
TASK AND SCHOOL ACHIEVEMENT
(BERQUAM, 1981)

APPENDIX N

CORRELATION COEFFICIENTS BETWEEN A SHORT TERM MEMORY TASK AND SCHOOL ACHIEVEMENT (BERQUAM, 1981)

	RO	R5	ACC	
CTBS	+.29	+.198	-.158	(n=38)

CTBS = Comprehensive Test of Basic Skills (Total Raw Score)

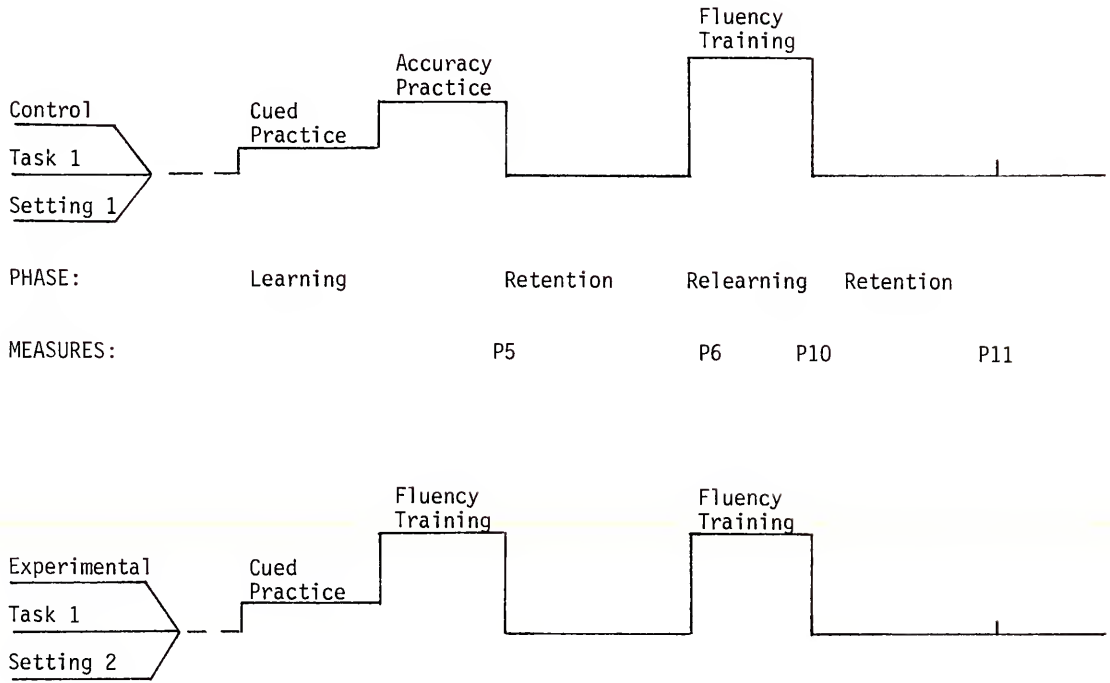
RO = Performance at end of learning phase (correct per minute)

R5 = Performance at end of retention phase (correct per minute)

ACC = performance at end of retention phase (percent correct)

APPENDIX O

EXPERIMENTAL DESIGN COMPONENTS



APPENDIX O

DESIGN COMPONENTS

APPENDIX P

MATCHING COMPARISONS

APPENDIX P

MATCHING COMPARISONS

	Mean	F-ratio	p
CTBS Scores*			
Total Reading			
Experimental	339.8		
Control	326.9		
		.377	.5453
Total Math			
Experimental	352.5		
Control	318.5		
		7.242	.0128
Total Battery			
Experimental	324.9		
Control	301.7		
		1.426	.2441
State Student Assessment Test Scores**			
Communication Skills			
Experimental	71.2		
Control	68.9		
		.649	.4279
Mathematics Skills			
Experimental	63.8		
Control	61.9		
		.750	.3945
Digit Writing Rate***			
Experimental	81.9		
Control	79.5		
		.141	.7089
Mean Correct Responses Per Student Per Day, During Learning Phase			
Experimental	9.7		
Control	9.6		

*Standard Score Units
**Raw Score Units
***Frequency (number correct per minute)

APPENDIX Q

CORRELATION COEFFICIENTS BETWEEN PERFORMANCE ON THE EXPERIMENTAL TASK AND SCHOOL ACHIEVEMENT

APPENDIX Q

CORRELATION COEFFICIENTS BETWEEN PERFORMANCE ON THE EXPERIMENTAL TASK AND SCHOOL ACHIEVEMENT

		SSAT		CTBS		
	DIGIT RATE	R	M	R	M	T
P5	.483	.326	.088	-.023	.258	.070
P6	.424	.263	-.032	-.223	-.006	-.114
P10	.367	.355	.071	-.018	.149	.078
P11	.348	.463	.167	.221	.171	.241

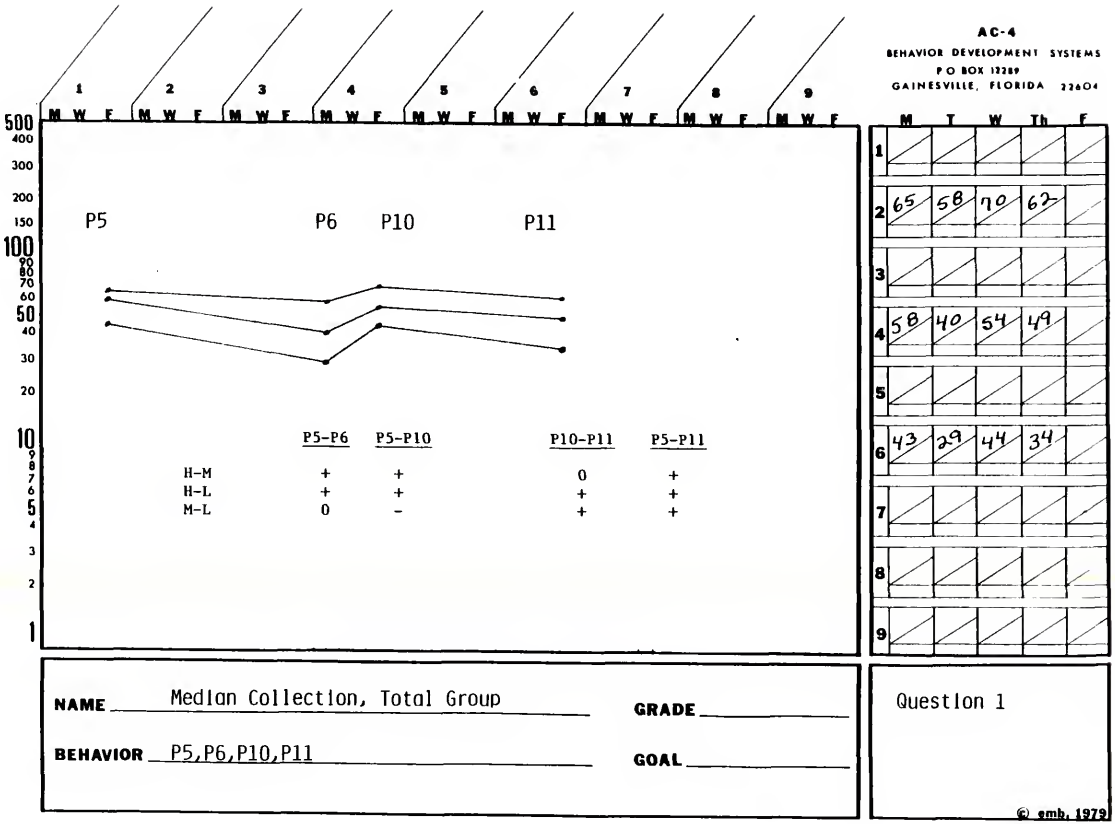
(N = 28)

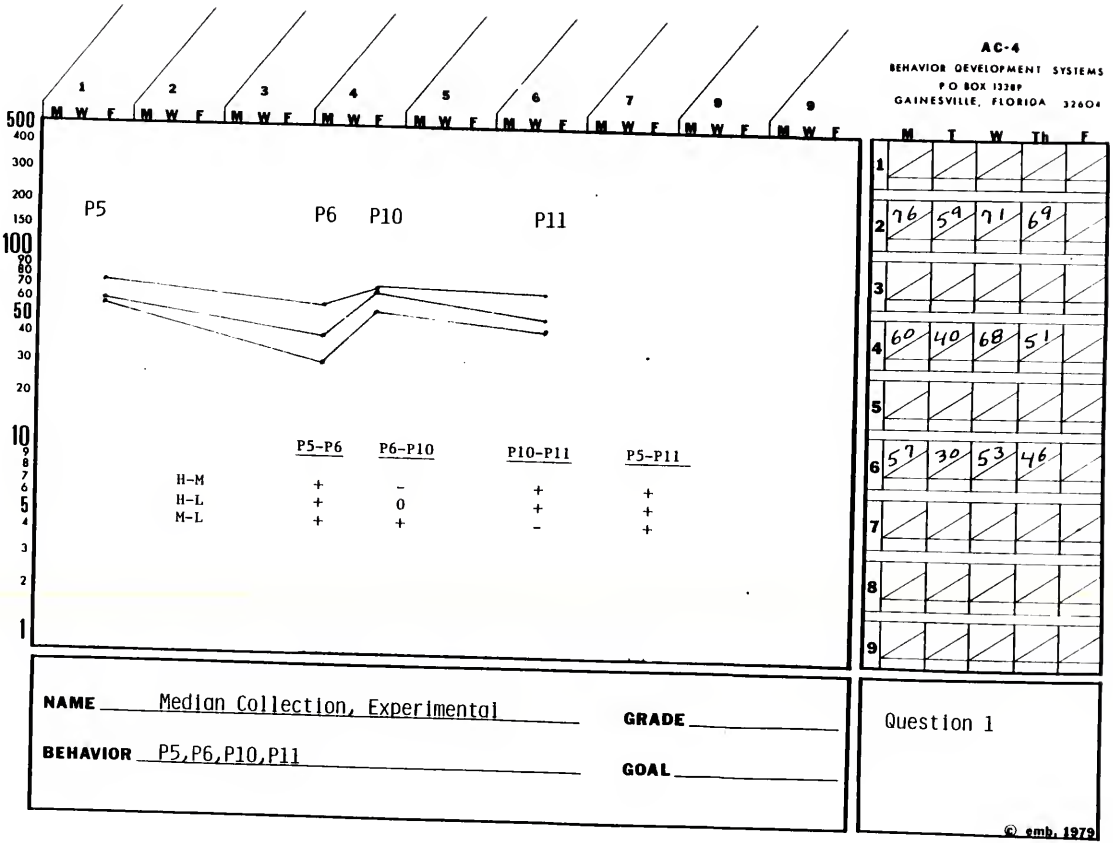
Variables

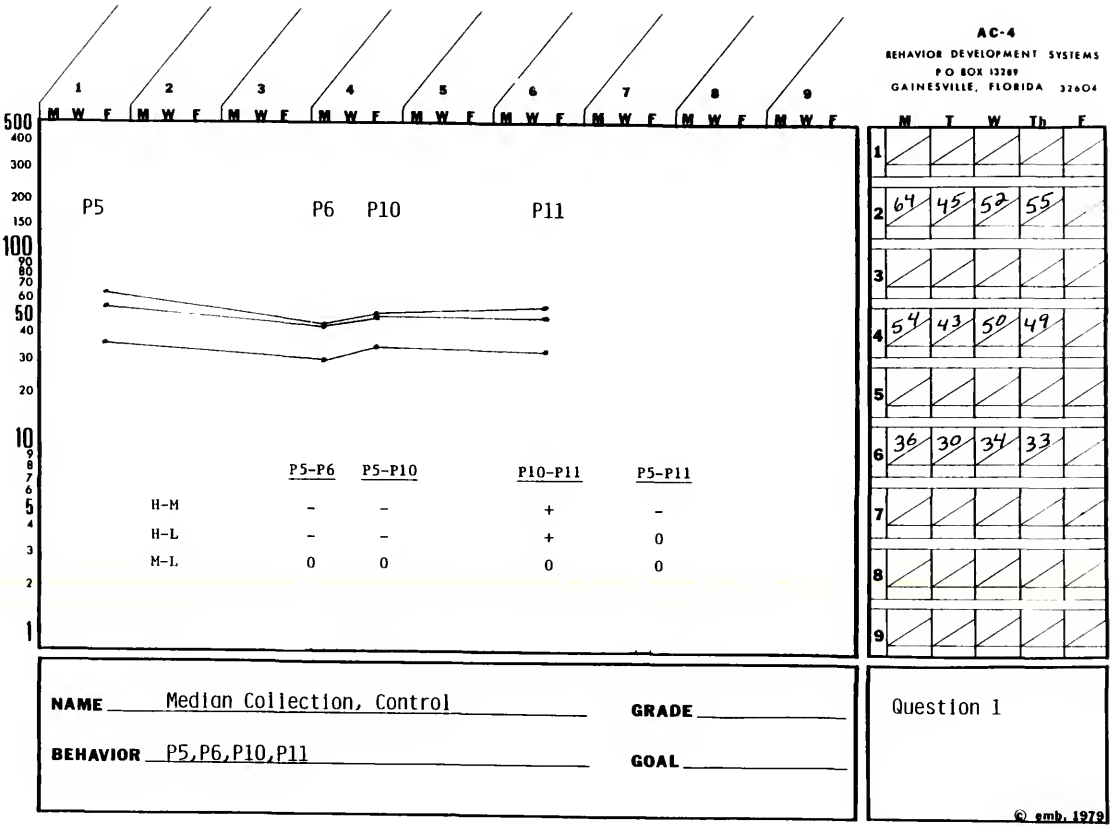
CTBS-R = Reading Subtest of Comprehensive Test of Basic Skills
CTBS-M = Math Subtest of Comprehensive Test of Basic Skills
CTBS-T = Total Raw Score on Comprehensive Test of Basic Skills
SSAT-R = Communications Subtest of Florida State Student Assessement Test
SSAT-M = Math Subtest of Florida State Student Assessment Test
P5 = Performance on Experimental Task, end of learning phase 1
P6 = Performance on Experimental Task, end of retention phase 1
P10 = Performance on Experimental Task, end of relearning phase
P11 = Performance on Experimental Task, end of retention phase 2

APPENDIX R

CHARTED DATA, QUESTION 1

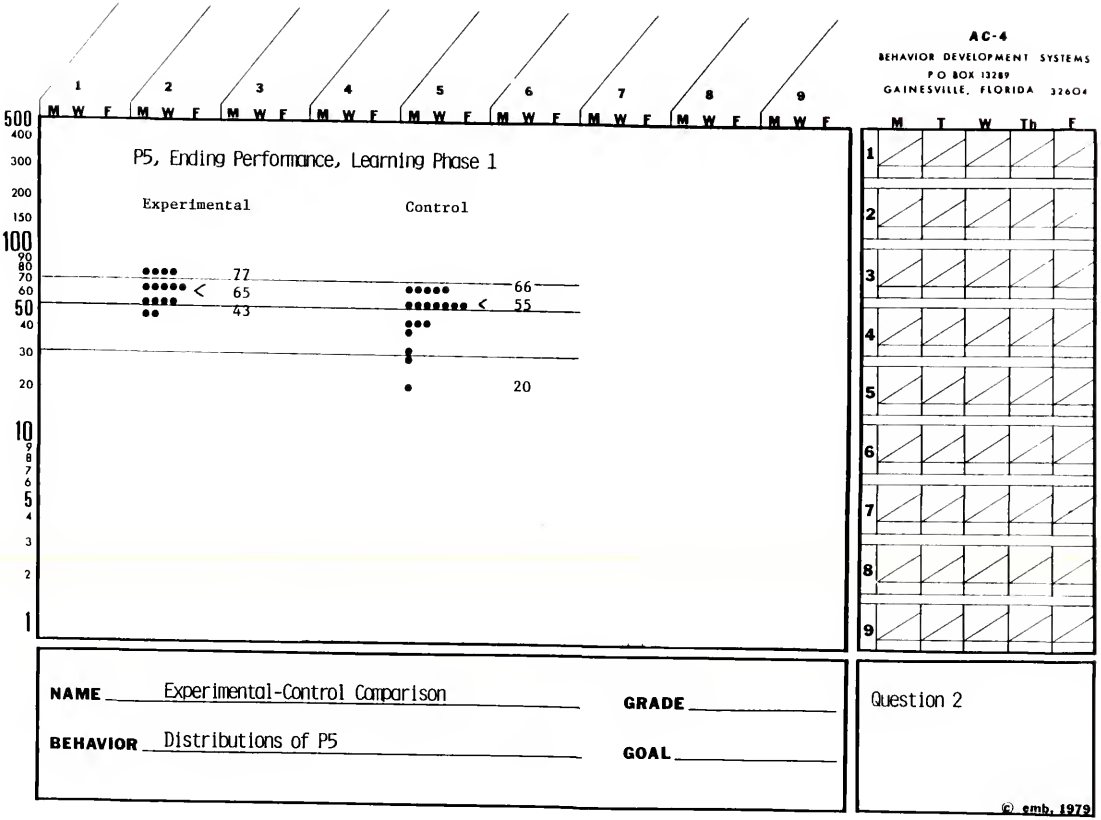

AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P O BOX 13289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500 400 300 200 150 100 90 80 70 60 50 40 30 20 10 9 8 7 6 5 4 3 2 1
P5 P6 P10 P11
P5-P6 P5-P10 P10-P11 P5-P11
H-M + + 0 +
H-L + + + +
M-L 0 - + +
M T W Th F
1
2 65 58 70 62
3
4 58 40 54 49
5
6 43 29 44 34
7
8
9
NAME Median Collection, Total Group
GRADE
BEHAVIOR P5,P6,P10,P11
GOAL
Question 1
© emh, 1979

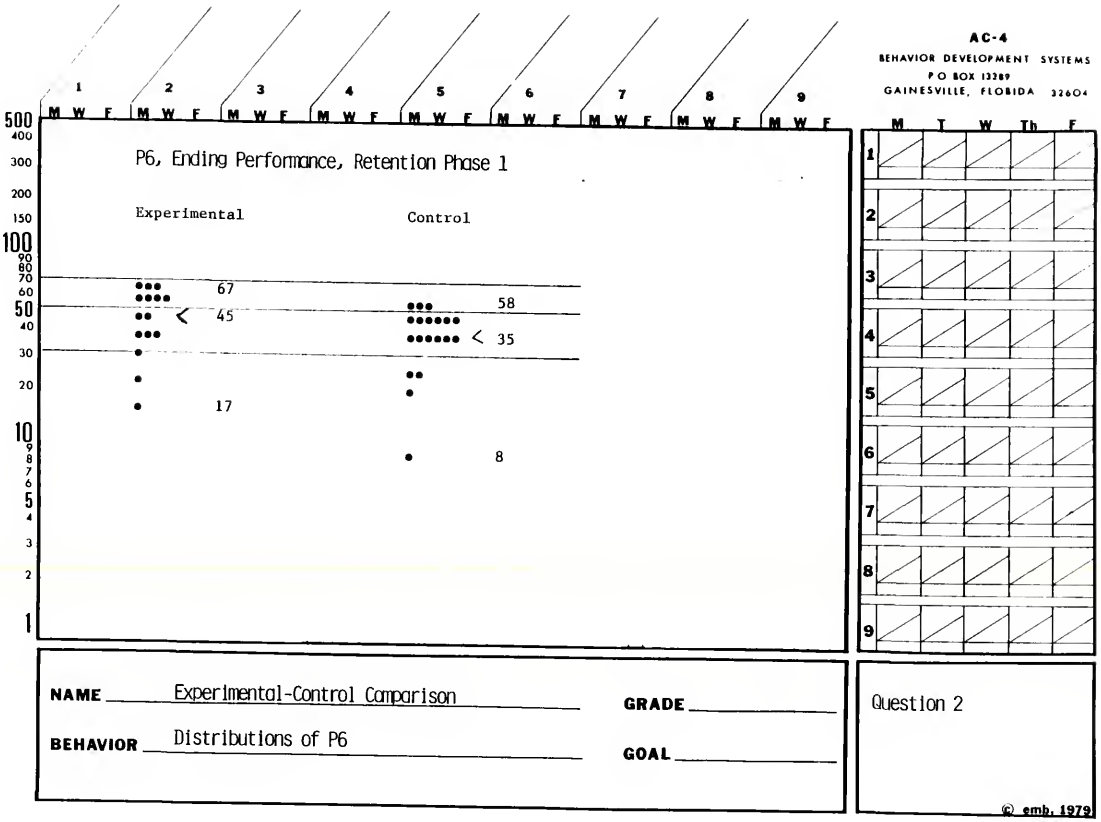

AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P O BOX 13289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500 400 300 200 150 100 90 80 70 60 50 40 30 20 10 9 8 7 6 5 4 3 2 1
P5 P6 P10 P11
P5-P6 P6-P10 P10-P11 P5-P11
H-M + - + +
H-L + 0 + +
M-L + + - +
M T W Th F
1
2 76 59 71 69
3
4 60 40 68 51
5
6 57 30 53 46
7
8
9
NAME Median Collection, Experimental
GRADE
BEHAVIOR P5,P6,P10,P11
GOAL
Question 1
© emb, 1979

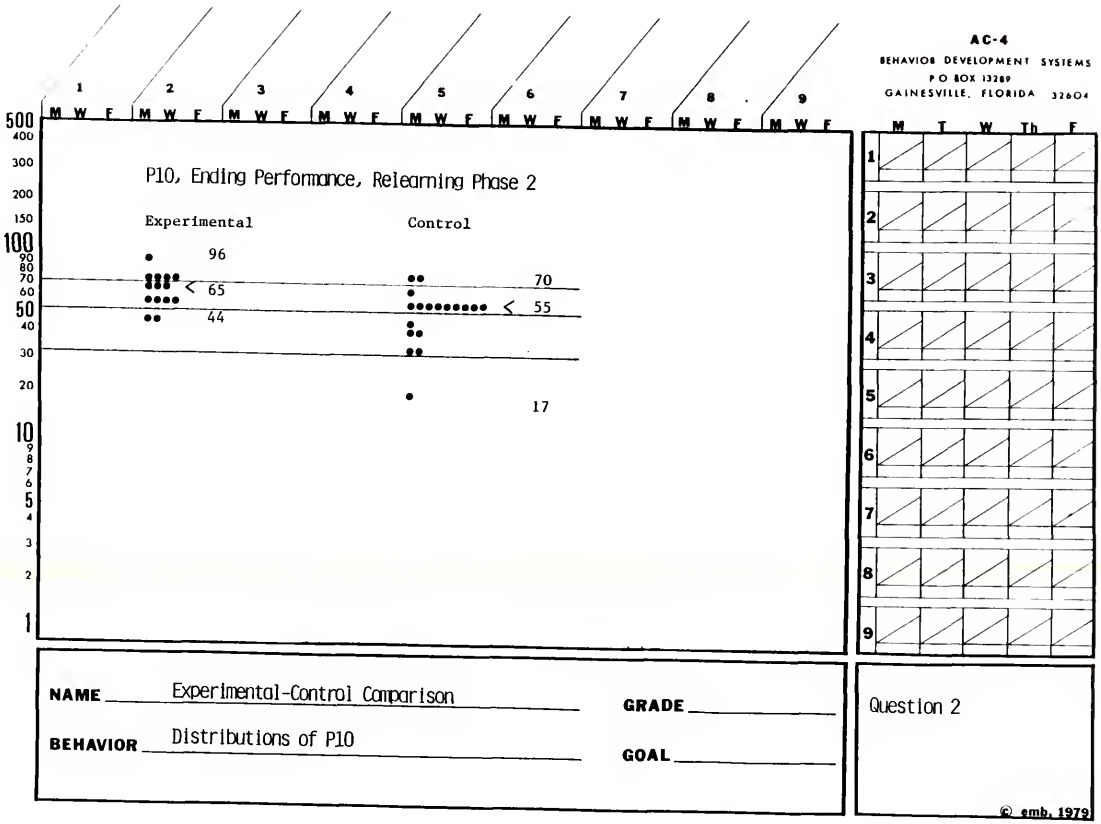

AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P.O. BOX 13289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500 400 300 200 150 100 90 80 70 60 50 40 30 20 10 9 8 7 6 5 4 3 2 1
P5 P6 P10 P11
P5-P6 P5-P10 P10-P11 P5-P11
H-M - - + -
H-L - - + 0
M-L 0 0 0 0
M T W Th F
1
2 64 45 52 55
3
4 54 43 50 49
5
6 36 30 34 33
7
8
9
NAME Median Collection, Control
GRADE
BEHAVIOR P5,P6,P10,P11
GOAL
Question 1
© emb, 1979

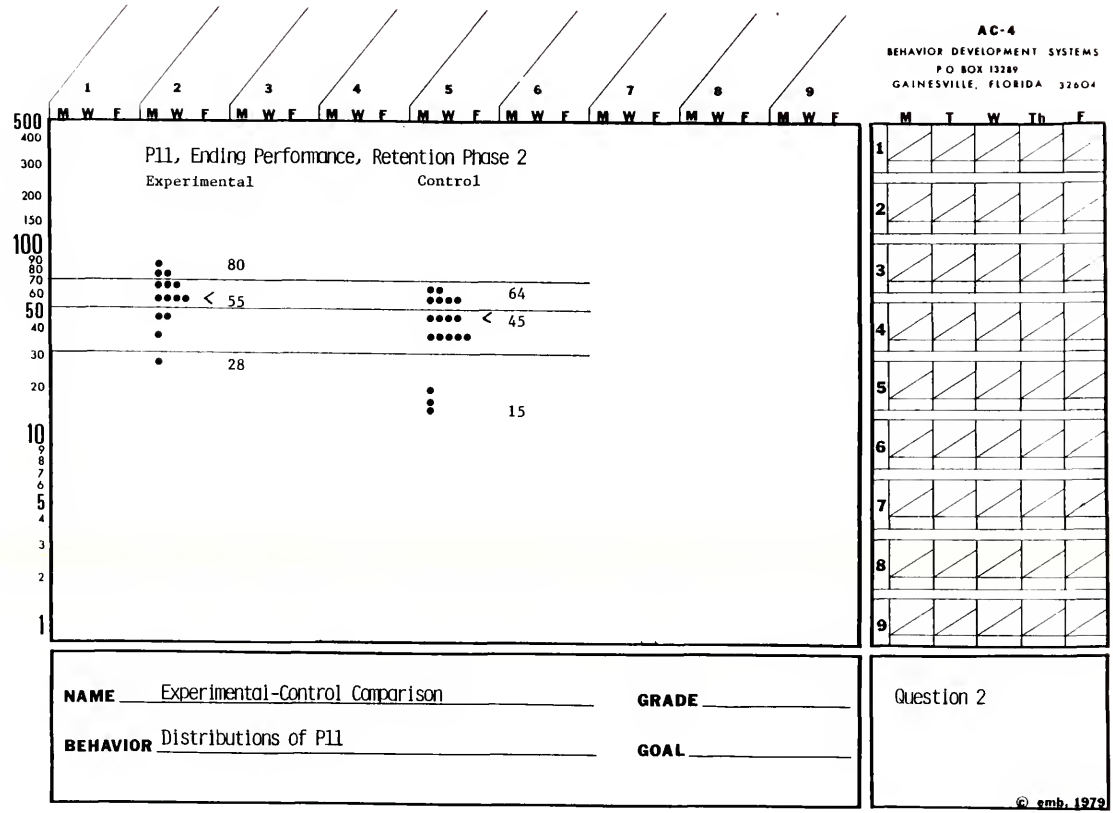
APPENDIX S

CHARTED DATA, QUESTION 2, FREQUENCY DISTRIBUTIONS


AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P.O. BOX 13289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500 400 300 200 150 100 90 80 70 60 50 40 30 20 10 9 8 7 6 5 4 3 2 1
P5, Ending Performance, Learning Phase 1
Experimental
Control
77
65
43
66
55
20
M T W Th F
1 2 3 4 5 6 7 8 9
NAME Experimental-Control Comparison
GRADE
BEHAVIOR Distributions of P5
GOAL
Question 2
© emb, 1979

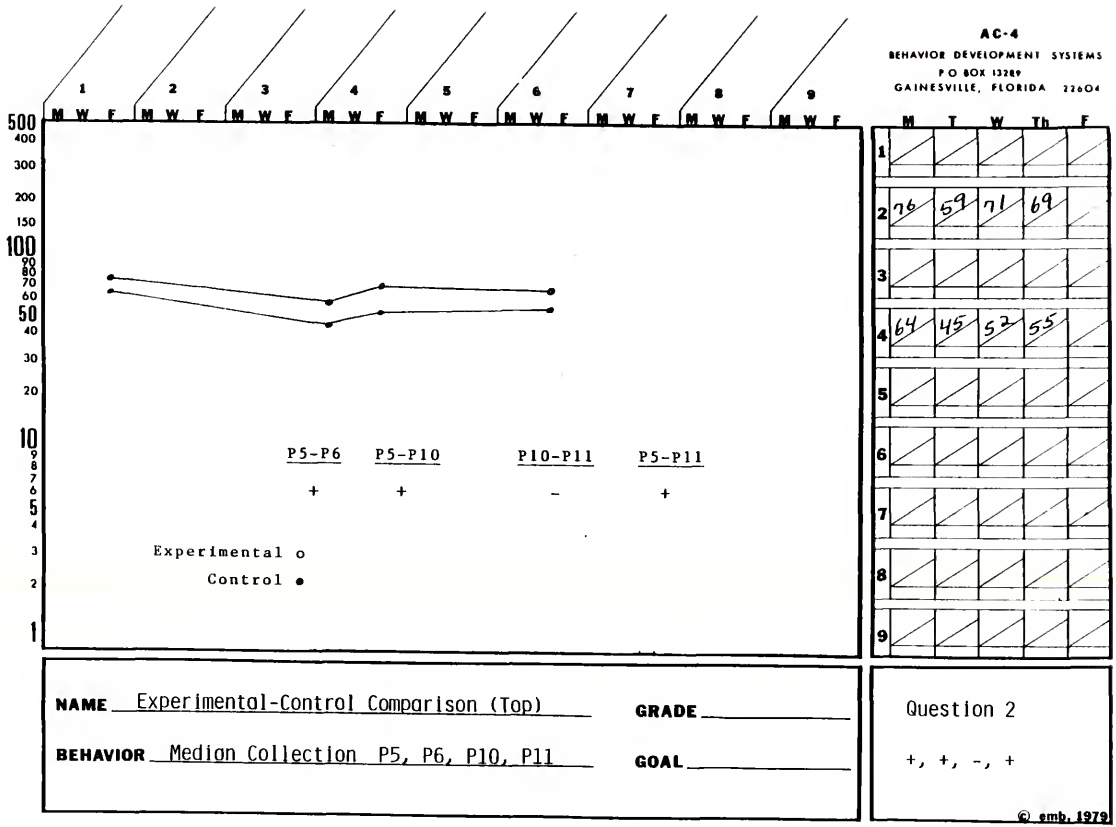

AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P O BOX 13289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
M T W Th F
500 400 300 200 150 100 90 80 70 60 50 40 30 20 10 9 8 7 6 5 4 3 2 1
P6, Ending Performance, Retention Phase 1
Experimental
Control
67
45
17
58
35
8
NAME Experimental-Control Comparison
GRADE
BEHAVIOR Distributions of P6
GOAL
Question 2
© emb, 1979

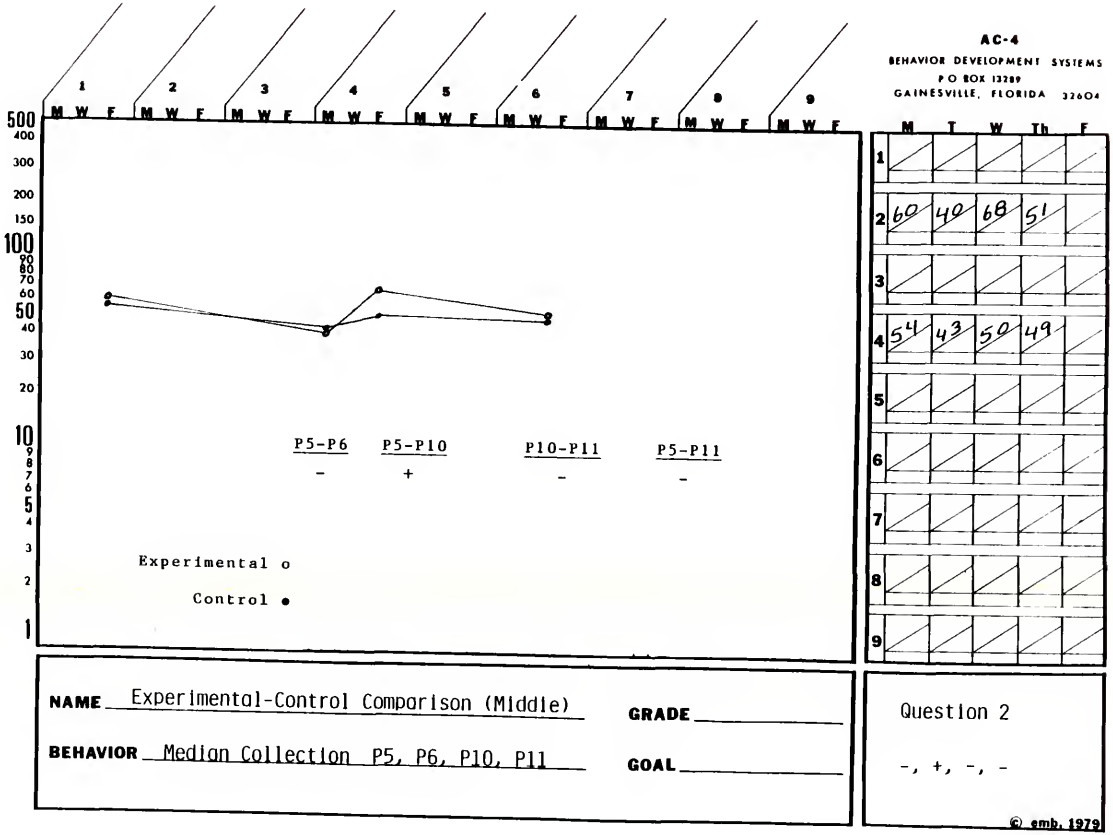

AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P O BOX 13289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
M T W Th F
500 400 300 200 150 100 90 80 70 60 50 40 30 20 10 9 8 7 6 5 4 3 2 1
P10, Ending Performance, Relearning Phase 2
Experimental
Control
96
65
44
70
55
17
NAME Experimental-Control Comparison
GRADE
BEHAVIOR Distributions of P10
GOAL
Question 2
© emb, 1979

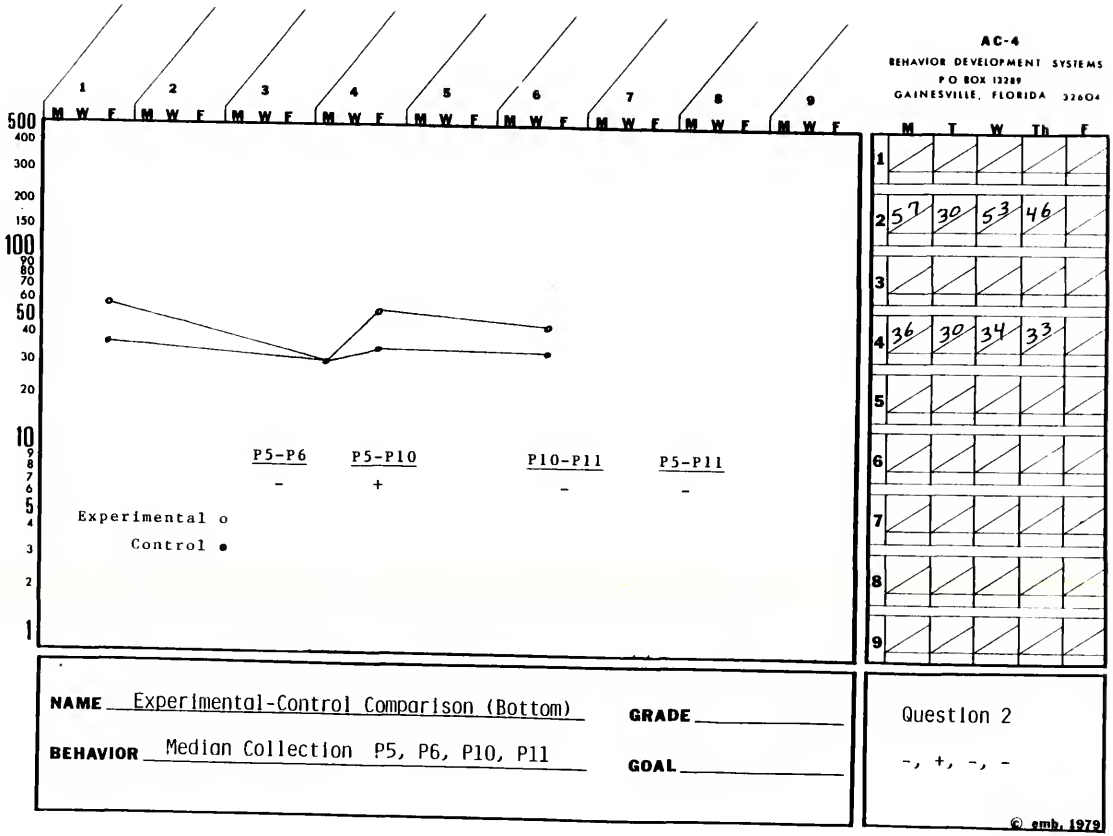

AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P.O. BOX 13289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500 400 300 200 150 100 90 80 70 60 50 40 30 20 10 9 8 7 6 5 4 3 2 1
P11, Ending Performance, Retention Phase 2
Experimental
Control
80
55
28
64
45
15
M T W Th F
1 2 3 4 5 6 7 8 9
NAME Experimental-Control Comparison
GRADE
BEHAVIOR Distributions of P11
GOAL
Question 2
© emb, 1979

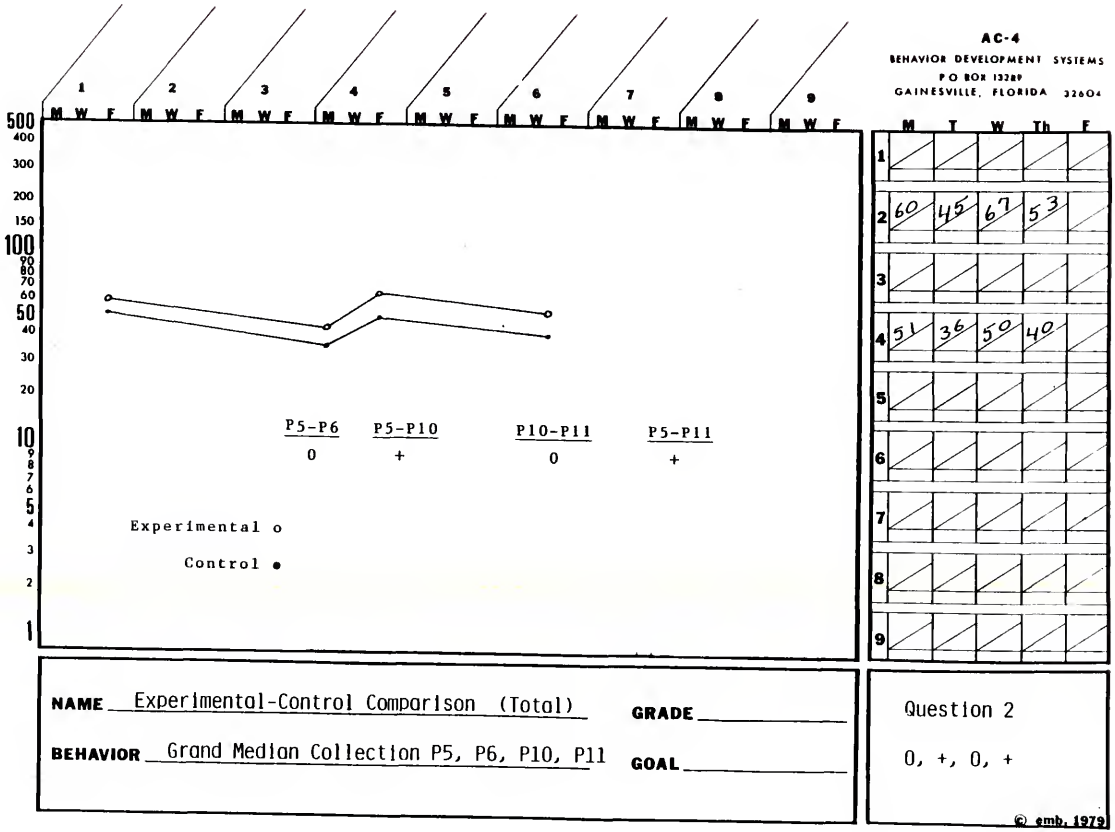
APPENDIX T

CHARTED DATA, QUESTION 2; EXPERIMENTAL-CONTROL COMPARISONS


AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P O BOX 13289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500 400 300 200 150 100 90 80 70 60 50 40 30 20 10 9 8 7 6 5 4 3 2 1
P5-P6 P5-P10 P10-P11 P5-P11
+ + - +
Experimental o
Control •
M T W Th F
1
2 76 69 71 69
3
4 64 45 52 55
5
6
7
8
9
NAME Experimental-Control Comparison (Top)
GRADE
BEHAVIOR Median Collection P5, P6, P10, P11
GOAL
Question 2
+, +, -, +
© emb, 1979

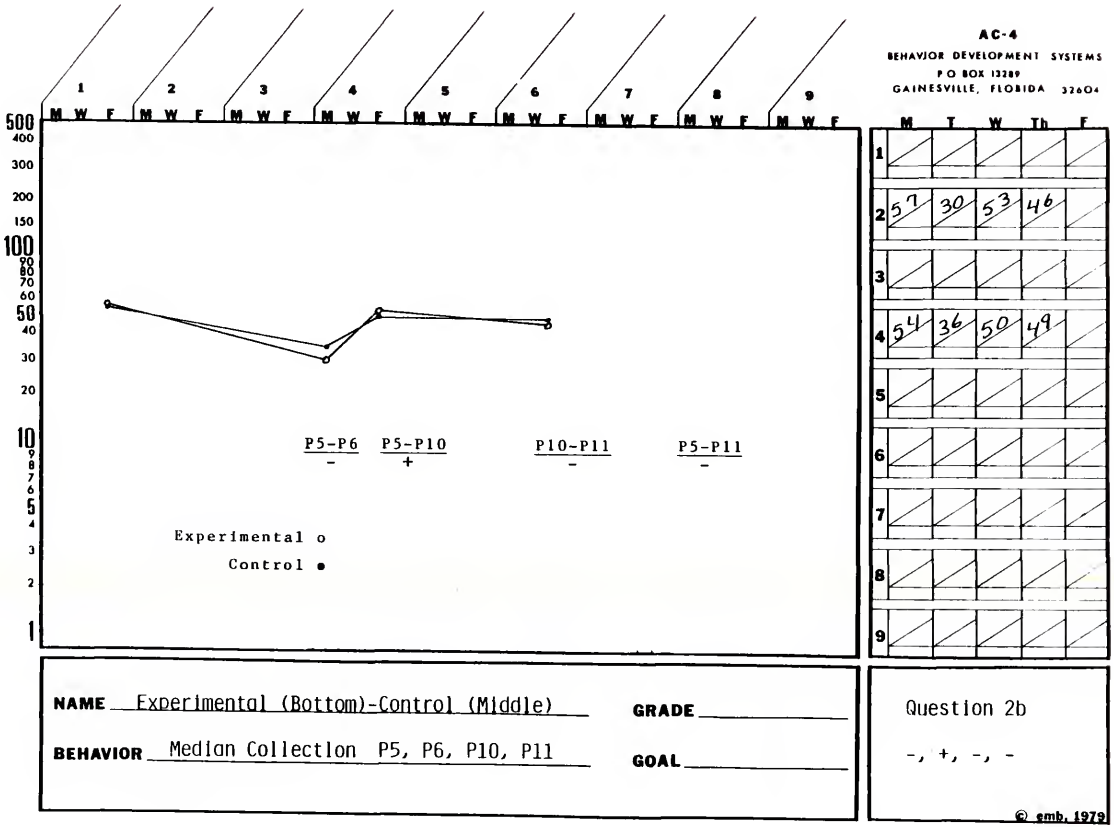

AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P O BOX 13289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500 400 300 200 150 100 90 80 70 60 50 40 30 20 10 9 8 7 6 5 4 3 2 1
P5-P6 P5-P10 P10-P11 P5-P11
- + - -
Experimental o
Control •
M T W Th F
1
2 60 40 68 51
3
4 54 43 50 49
5
6
7
8
9
NAME Experimental-Control Comparison (Middle)
GRADE
BEHAVIOR Median Collection P5, P6, P10, P11
GOAL
Question 2
-, +, -, -
© emb, 1979

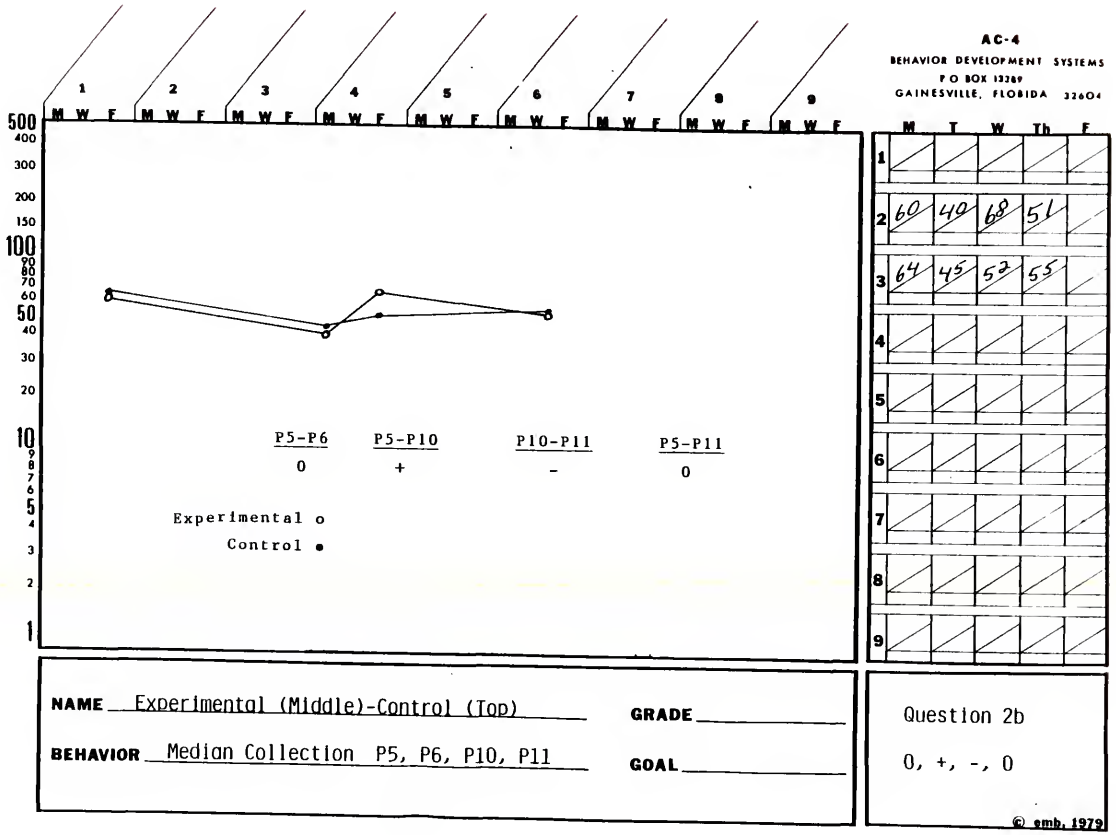

AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P O BOX 13289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500 400 300 200 150 100 90 80 70 60 50 40 30 20 10 9 8 7 6 5 4 3 2 1
P5-P6 P5-P10 P10-P11 P5-P11
- + - -
Experimental o
Control •
M T W Th F
1
2 57 30 53 46
3
4 36 30 34 33
5
6
7
8
9
NAME Experimental-Control Comparison (Bottom)
GRADE
BEHAVIOR Median Collection P5, P6, P10, P11
GOAL
Question 2
-, +, -, -
© emb, 1979


AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P.O. BOX 13289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500 400 300 200 150 100 90 80 70 60 50 40 30 20 10 9 8 7 6 5 4 3 2 1
M T W Th F
1
2 60 45 67 53
3
4 51 36 50 40
5
6
7
8
9
P5-P6 P5-P10 P10-P11 P5-P11
0 + 0 +
Experimental o
Control •
NAME Experimental-Control Comparison (Total)
GRADE
BEHAVIOR Grand Median Collection P5, P6, P10, P11
GOAL
Question 2
0, +, 0, +
© emb, 1979

APPENDIX U

CHARTED DATA, QUESTION 2; EXPERIMENTAL-CONTROL COMPARISONS, MATCHED ON P5

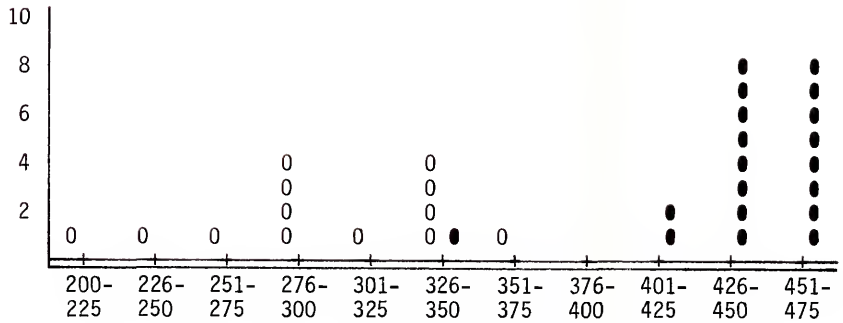

AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P.O. BOX 13289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500 400 300 200 150 100 90 80 70 60 50 40 30 20 10 9 8 7 6 5 4 3 2 1
P5-P6 -
P5-P10 +
P10-P11 -
P5-P11 -
Experimental o
Control •
M T W Th F
1
2 57 30 53 46
3
4 54 36 50 49
5
6
7
8
9
NAME Experimental (Bottom)-Control (Middle)
GRADE
BEHAVIOR Median Collection P5, P6, P10, P11
GOAL
Question 2b
-, +, -, -
© emb, 1979


AC-4
BEHAVIOR DEVELOPMENT SYSTEMS
P O BOX 13289
GAINESVILLE, FLORIDA 32604
1 2 3 4 5 6 7 8 9
M W F
500 400 300 200 150 100 90 80 70 60 50 40 30 20 10 9 8 7 6 5 4 3 2 1
P5-P6 0
P5-P10 +
P10-P11 -
P5-P11 0
Experimental o
Control •
M T W Th F
1
2 60 40 68 51
3 64 45 57 55
4
5
6
7
8
9
NAME Experimental (Middle)-Control (Top)
GRADE
BEHAVIOR Median Collection P5, P6, P10, P11
GOAL
Question 2b
0, +, -, 0
© emb, 1979

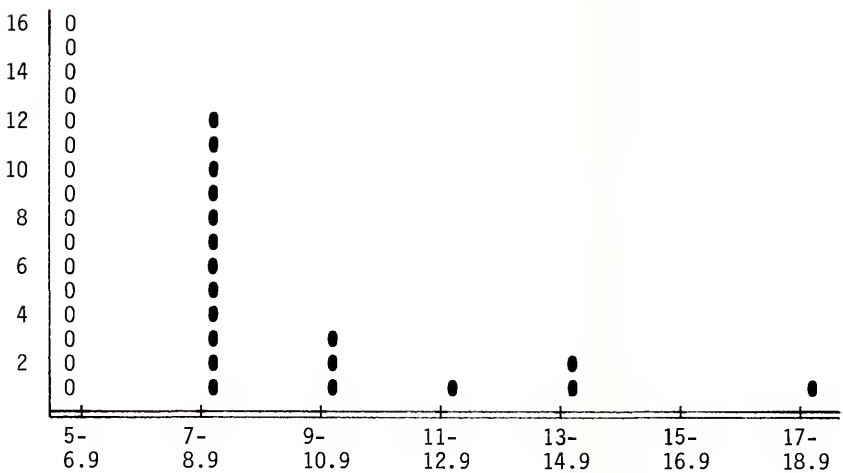
APPENDIX V

CHARTED DATA, QUESTION 2; DISTRIBUTIONS OF TOTAL CORRECT RESPONSES AND TOTAL TIME ON TASK

APPENDIX V



Total Correct Responses During Learning Phase



Total Minutes on Task During Learning Phase

O=Experimental ●=Control

REFERENCES

Ammons, R. B., Farr, R. G., Bloch, E., Neumann, E., Dey, M., Marion, R., Ammons, C. Long term retention of perceptual motor skills. Journal of Experimental Psychology, 1958, 55, 318-328.

Berquam, E. M. Academic behavior chart (AC-4). Gainesville, FL: Behavior Development Systems, 1979.

Berquam, E.M. Frequency of response and retention: Preliminary work. (Unpublished working paper, 1981).

Berquam, E.M. STRAT analysis. Journal of Precision Teaching (in press).

Brehmer, B. & Lindberg, L. Retention of single cue probability learning tasks as a function of cue validity, retention interval, and degree of learning. Journal of Experimental Psychology, 1973, 101, 406-408.

Briggs, G. E. Retroactive inhibition as a function of the degree of original and interpolated learning. Journal of Experimental Psychology, 1957, 53, 60-67.

Buxton, C. E. Level of mastery and reminiscence in pursuit learning. Journal of Experimental Psychology, 1943, 32, 176-180.

Carroll, J.B. Problems of measurement related to the learning for mastery. Educational Horizons, 1970, 48, 71-80.

Chasey, W. C. Overlearning as a variable in the retention of gross motor skills by the mentally retarded. Research Quarterly, 1971, 42, 145-149.

Chasey, W. C. Motor skill overlearning effects on retention and relearning by retarded boys. Research Quarterly, 1977, 48, 41-46.

Chasey, W. C. & Knowles, C. J. Effects of overlearning on retention and relearning of gross-motor skill by mentally retarded males. Perceptual and Motor Skills, 1973, 36, 503-509.

Clement, R. Spring-fall retention scores on precise monthly assessments. Great Falls Public Schools, Great Falls, Montana, 1978.

Ebbinghaus, H. Memory: A Contribution to experimental psychology. (H. A. Ruger & C. E. Bussenius, Translation.) New York: Teachers College, Columbia University, 1913. (Originally published, 1885.)

Estes, W. K., Hopkins, B. L. & Crothers, E. J. All or none and conservation effects in the learning and retention of paired associates. Journal of Experimental Psychology, 1960, 60, 329-339.

Fleishman, E. A., & Parker, J. F. Factors in the retention and relearning of perceptual-motor skills. Journal of Experimental Psychology, 1962, 64, 215-226.

Frahm, R. & Covington, J. What's Happening in Minimum Competency Testing? Bloomington, Indiana: Phi Delta Kappa, 1979.

Gagne, R. M. The conditions of learning. New York: Holt, Rinehart and Winston, Inc., 1965.

Gilbert, T. F. Overlearning and the retention of meaningful prose. Journal of General Psychology, 1957, 56, 281-289.

Gillette, A. L. Learning and retention: A comparison of three experimental procedures. Archives of Psychology, 1936, 28, No.198.

Hall, J.F. The psychology of learning. New York: J.B. Lippincott Company, 1966.

Hammerton, M. Retention of learning in a difficult tracking task. Journal of Experimental Psychology, 1963, 66, 108-110.

Haring, N.G., Lovitt, T. C., Eaton, M. D. & Hansen, C. The fourth r: Research in the classroom. Columbus, Ohio: Charles E. Merrill, 1978.

Hellyer, S. Frequency of stimulus presentation and short-term decrement of recall. Journal of Experimental Psychology, 1962, 55, 270-273.

Hilgard, E. R. The savings score as a measure of retention. American Journal of Psychology, 1934, 46, 337-339.

Hilgard, E. R. Methods and procedures in the study of learning. In S. S. Stevens (editor) Handbook of experimental psychology. New York: John Wiley & Son, 1951.

Hovland, C. I. Human learning and retention. In S. S. Stevens (editor) Handbook of experimental psychology. New York: John Wiley & Son, 1951.

Hunter, C.S., & Harman, D. Adult illiteracy in the United States. New York: Ford Foundation, 1979.

Irion, A. L. Reminiscence in pursuit-rotor learning as a function of length of rest and amount of prerest practice. Journal of Experimental Psychology, 1949, 39, 492-499.

Jahnke, J. C. Retention in motor learning as a function of amount of practice and rest. Journal of Experimental Psychology, 1958, 55, 270-273.

Johnston, J. M., & Pennypacker, H. S. Strategies and tactics of human behavioral research. Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc., 1980.

Kline, C. L. Adolescents with learning problems: How long must they wait? Journal of Learning Disabilities, 1972, 5, 262284.

Kratochwill, T., Demuth, D., & Conzemius, W. Effects of overlearning on preschool children's retention of sight word vocabulary words. Reading Improvement, 1977, 14, 223-226.

Krueger, W. C. The effect of overlearning on retention. Journal of Experimental Psychology, 1929, 12, 71-79.

Krueger, W. C. Further studies in overlearning. Journal of Experimental Psychology, 1930, 13, 152-163.

Lance, W. D. Effects of meaningfulness and overlearning on retention in normal and retarded adolescents. American Journal of Mental Deficiency, 1965, 70, 270-275.

Luh, C. W. The conditions of retention. Psychological Monographs, 1922, 31, No. 3.

McGeoch, J. A. The influence of degree of learning upon retroactive inhibition. American Journal of Psychology, 1929, 41, 252-262.

McGeoch, J. A., & Irion, A. The psychology of human learning. New York: David McKay Company, Inc., 1965.

Melnick, M. J. Effects of overlearning on the retention of a gross-motor skill. Research Quarterly, 1971, 42, 60-69.

Mordock, J. B. Paired associate learning in mental retardation: A review. American Journal of Mental Deficiency, 1968, 72, 857-865.

Pennypacker, H. S., Koenig, C. H., & Lindsley, O. R. Handbook of the standard behavior chart. Kansas City, Kansas: Precision Media, 1972.

Peterson, L. R., & Peterson, M. J. Short-term retention of individual verbal items. Journal of Experimental Psychology, 1959, 58, 193-198.

Postman, L. Retention as a function of degree of overlearning. Science, 1962, 135, 666-667.

Purdy, B. J., & Lockhart, A. Retention and relearning of gross motor skills after long periods of no practice. Research Quarterly, 1962, 33, 265-272.

Reynolds, T., & Glaser, R. Effects of repetition and spaced review. Journal of Educational Psychology, 1964, 55, 297-308.

Richardson, J. Retention of concepts as a function of the degree of original and interpolated learning. Journal of Experimental Psychology, 1956, 51, 358-364.

Roberts, W. A. Short-term memory in the pigeon: Effects of repetition and spacing. Journal of Experimental Psychology, 1972, 94, 74-83.

Rock, I. Repetition and learning. Scientific American, 1958, 199, 68-72.

Rubin-Rabson, G. Studies in the psychology of memorizing piano music. VI: A comparison of two forms of mental rehearsal and keyboard overlearning. Journal of Educational Psychology, 1941, 32, 593-601. (a)

Rubin-Rabson, G. Studies in the psychology of memorizing piano music. VII: A comparison of three degrees of overlearning. Journal of Educational Psychology, 1941, 32, 688-696. (b)

Schmid, C. F. Handbook of graphic presentation. Seattle: Ronald Press Co., 1954.

Sidman, M. Tactics of scientific research. New York: Basic Books, Inc., 1960.

Skinner, B. F. Some contributions of an experimental analysis of behavior to psychology as a whole. American Psychologist, 1953, 8, 69-78.

Slamecka, T. Retention of connected discourse. American Journal of Psychology, 1959, 72, 409-416.

Spicker, H.H. Selected factors influencing learning and retention abilities of the mentally retarded. Education and Training of the Mentally Retarded, 1966, 1, 92-94.

Underwood, B. J. Studies of distributive practice: XII. Retention following varying degrees of original learning. Journal of Experimental Psychology, 1954, 47, 294-300.

Underwood, B. J. Forgetting. Scientific American, 1964, 210, 91-99. (a)

Underwood, B. J. Degree of learning and the measurement of forgetting. Journal of Verbal Learning and Verbal Behavior, 1964, 3, 112-129. (b)

Underwood, B. J. & Keppel, G. Retention as a function of degree of learning and letter sequence interference. Psychological Mongraphs, 1963, 77, 1-16.

Vergason, G. A. Retention in retarded and normal subjects as a function of amount of original training. American Journal of Mental Deficiency, 1964, 68, 623-629.

Vergason, G. A. Facilitation of memory in the retardate. Exceptional Children, 1968, 34, 589-594.

Wellborn, S. N. Kids, teachers, parents: "Give us better schools." U.S. News & World Report, September 10, 1979, pp. 31-32.

White, O. R. The "split middle," a "quickie" method of trend estimation. Working paper, Experimental Education Unit, University of Washington, 1972.

White, O. R., & Haring, N. G. Exceptional Teaching. Columbus, Ohio: Charles E. Merrill, 1980.

Wolking, W. D. A primer of precision teaching. Gainesville, Florida: Behavior Development Systems, 1972.

Wolking, W. D. Precision educational measurement: The direct and continuous measurement of pupil behavior: A self-instruction module. Florida State Department of Education, Tallahassee, Florida, 1972.

Wolking, W. D., Ehrhardt, D. & Nancarrow, E. The Florida language profile. Baltimore: ERIC Document Reproduction Service, Inc., 1973.

Wollen, K. A. One trial versus incremental paired associate learning. Journal of Verbal Learning and Verbal Behavior, 1962, 1, 14-21.

Woodworth, R. S. Experimental psychology. New York: Holt, 1938.

BIOGRAPHICAL SKETCH

Eugene "Skip" Berquam was born in 1946, in Moorhead, Minnesota. After graduation from the local high school, he enrolled at Moorhead State University, and worked to squeeze four years of college into six, graduating in 1970 with a degree in elementary education.

Four years of teaching sixth grade in Jamestown, North Dakota, taught him two things. One, he liked working with kids, and two, he didn't know much about teaching. Convinced that he did not want to be a school administrator, he entered a graduate program in education at Western Illinois University, and graduated in 1974 as an elementary principal.

After a year of work as a teaching principal in Fort Smith, Montana, it became clear to Skip that he still did not know much about how kids learn. When he was first introduced to precision teaching, he immediately saw the potential for its use in education. In the next three years at Fort Smith, it was a toss-up as to who learned more, the students or Mr. Berquam. He became more deeply involved with learning about and developing and implementing precision teaching programs in his school.

In 1978, Skip enrolled at the University of Florida and began learning about the science of behavior. Since 1980, he has been administrator of a precision teaching project in Orange County, Florida, developing fluency-based basic skills materials.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William D. Wolking, Chairman
Professor of Special Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Robert A. Algozzine
Associate Professor of
Special Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Cecil D. Mercer
Professor of Special Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Ronald E. Nutter
Assistant Professor of
Special Education

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

H. S. Pennypacker
Professor of Psychology

This dissertation was submitted to the Graduate Faculty of Special Education in the College of Education and to the Graduate Council, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

June, 1981

Dean for Graduate Studies

UNIVERSITY OF FLORIDA
3 1262 08666 351 4